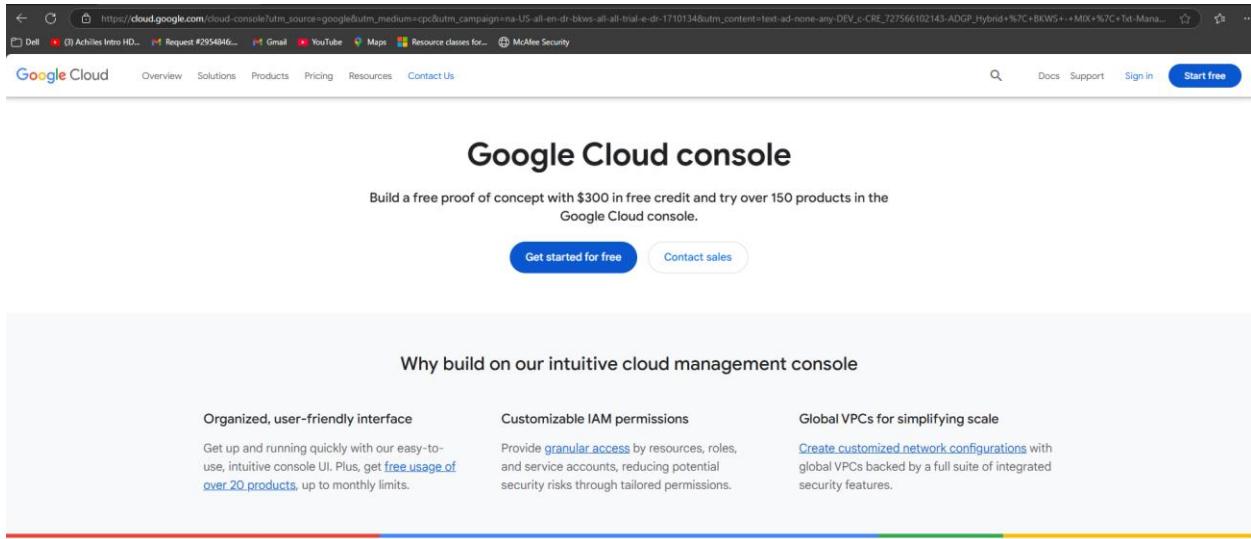


GCP Databricks End To End Project

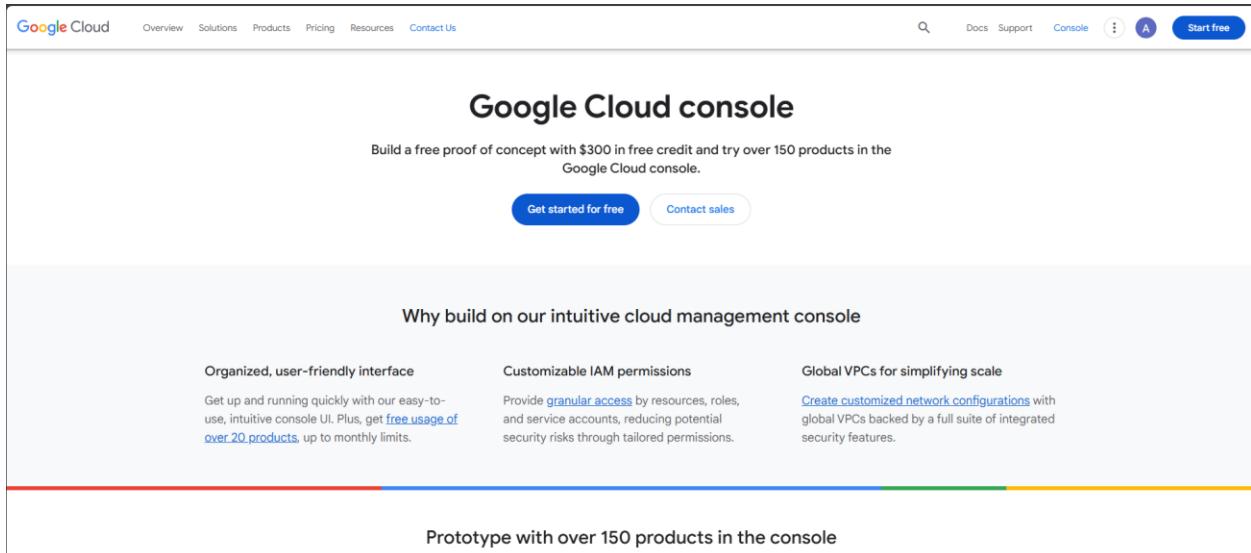
Create Google Cloud Account



The screenshot shows the Google Cloud console homepage. At the top, there's a navigation bar with links for Overview, Solutions, Products, Pricing, Resources, Contact Us, Docs, Support, Sign in, and a Start free button. Below the navigation is a search bar. The main heading is "Google Cloud console". A sub-headline says, "Build a free proof of concept with \$300 in free credit and try over 150 products in the Google Cloud console." There are two buttons: "Get started for free" and "Contact sales". Below this, a section titled "Why build on our intuitive cloud management console" lists three features: "Organized, user-friendly interface", "Customizable IAM permissions", and "Global VPCs for simplifying scale". Each feature has a brief description and a link.

Organized, user-friendly interface	Customizable IAM permissions	Global VPCs for simplifying scale
Get up and running quickly with our easy-to-use, intuitive console UI. Plus, get free usage of over 20 products , up to monthly limits.	Provide granular access by resources, roles, and service accounts, reducing potential security risks through tailored permissions.	Create customized network configurations with global VPCs backed by a full suite of integrated security features.

Login with your gmail account credentials



This screenshot is identical to the one above it, showing the Google Cloud console homepage. It includes the same navigation bar, search bar, main heading, sub-headline, and "Get started for free" and "Contact sales" buttons. The "Why build on our intuitive cloud management console" section is also present with its three listed features and descriptions.

https://console.cloud.google.com/freetrial/signup/tos?facet_utm_source=google&facet_utm_campaign=na-US-all-en-dr-bkws-all-all-trial-e-dr-1710134&facet_utm_medium=cpc&facet_url=https%2F%2Fcloud.google.com/freetrial

Dell (3) Achilles Intro HD... Request #2954846... Gmail YouTube Maps Resource classes for... McAfee Security

Try Google Cloud for free

Step 1 of 2 Account Information



Abhijit Das

[REDACTED]@gmail.com

[Switch account](#)

Country

United States

By using this application, you agree to the [Google Cloud Platform Terms of Service](#), [Supplemental Free Trial Terms of Service](#), and [any applicable services and APIs Terms of Service](#).

[Agree & continue](#)

Access to all Google Cloud products

Get everything you need to build and run your apps, websites and services, including Firebase and the Google Maps API.

\$300 in free credit

Try Google Cloud with \$300 in credit to spend over the next 90 days.

No automatic charges

You only start paying if you decide to activate a full, pay-as-you-go account or choose to prepay. You'll keep any remaining free credit.

https://console.cloud.google.com/freetrial/signup/billing/US?facet_utm_source=google&facet_utm_campaign=na-US-all-en-dr-bkws-all-all-trial-e-dr-1710134&facet_utm_medium=cpc&facet_url=https%2F%2Fcloud.google.com/freetrial

Dell (3) Achilles Intro HD... Request #2954846... Gmail YouTube Maps Resource classes for... McAfee Security

Try Google Cloud for free

Step 2 of 2 Verify Your Identity

We use payment information to verify your identity, which prevents abuse of our services by spammers and bots. **You won't be charged unless you activate your full, pay-as-you-go account or choose to prepay.**

Payments profile

Abhijit Das
Individual • Also used with Google Pay • United States • [Change](#)

ID [REDACTED]
Full address required

Your payment information is saved in a payments profile, which is associated with your Google Account and shared across Google services. [Learn more about payments profile](#)

Payment method

[REDACTED] Visa [REDACTED] [Change](#)

Please complete the previous sections before continuing

[Start free](#)

Access to all Google Cloud products

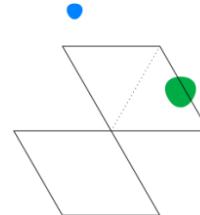
Get everything you need to build and run your apps, websites and services, including Firebase and the Google Maps API.

\$300 in free credit

Try Google Cloud with \$300 in credit to spend over the next 90 days.

No automatic charges

You only start paying if you decide to activate a full, pay-as-you-go account or choose to prepay. You'll keep any remaining free credit.



[Privacy policy](#) | [FAQs](#)

Google Cloud Account (Billing tab)

The screenshot shows the Google Cloud Billing Overview page. It displays a summary of costs: Cost \$0.00, Credits used \$0.00, and Total cost \$0.00. A sidebar on the left provides links for reports, cost breakdown, budgets, and anomalies. On the right, there's a section for setting a monthly budget with a 'Create' button, a 'Billing health checks' section with one critical issue, and a 'Free trial credit' section showing \$300 available. A 'Recommended for you' sidebar on the right lists various tutorials and documents related to Cloud Billing.

Create a separate project for implementing Databricks Project.

The screenshot shows the 'New Project' creation page. It asks for a project name (GCP Databricks Project), which cannot be changed later. It also asks for a location (No organization) and a parent organization or folder. A note says there are 11 projects remaining in quota. Buttons for 'Create' and 'Cancel' are at the bottom.

The screenshot shows the 'Welcome' screen for the new project 'My First Project'. It displays a message 'Welcome, Abhijit Das', a summary of being in a free trial (0 out of \$300 credits used, expires July 3, 2025), and options to activate full account, add people, set up budget alerts, and review product spend. A notifications sidebar on the right shows two recent notifications about project creation.

To create GCP Databricks account, we need to subscribe to Databricks.

Find Databricks under Marketplace (Partner Solutions)

Google Cloud My First Project

Filter Type to filter

Category

- Analytics (1,117)
- Big data (584)
- Business Applications (209)
- Contact Center AI (13)
- Databases (504)

Analytics

 commercetools Reimaging digital commerce for the world's leading companies	 FullStory FullStory, INC Digital Experience Intelligence and Product Analytics	 LiveRamp LiveRamp Improves marketing by collaborating, joining, activating and measuring data across...	 Quantum Metric Quantum Metric The Platform For Continuous Product Design	 Shopify Shopify Inc The leading global commerce company that provides essential internet infrastructure for...	 eCommerce Search, Browse, & Recommendations Groupby B2B & B2C composable eCommerce Product Discovery Platform powered by Google Clou...
--	---	--	---	---	--

Type

- Google Cloud Platform (36)
- SaaS & APIs (1,648)
- Virtual machines (1,919)
- Data (165)
- Vertex AI (23)

Price

Analytics

 Apache Kafka® & Apache Flink® on Confluent Cloud™ Confluent Apache Kafka, reinvented for the modern data stack	 Databricks Databricks All your data, analytics and AI on one Lakehouse platform	 Fivetran Data Pipelines Fivetran Automated data integration for zero-maintenance data pipelines	 Informatica Intelligent Data Management Cloud Informatica Better Data. Even Smarter Analytics.	 Teradata VantageCloud Enterprise on Google Cloud Teradata The connected multi-cloud data platform for enterprise analytics.
---	--	--	---	---

DevOps

 Datadog Datadog See inside any stack, any app, at any scale, anywhere	 GitLab GitLab GitLab is a comprehensive AI-powered DevSecOps platform	 Dynatrace SaaS Dynatrace LLC Unified observability & security for cloud and on-premises	 Snyk: Developer Security Platform Snyk Ltd. Snyk: Developer Security Platform	 Harness Continuous Delivery Harness, Inc. Continuous Delivery as a Service
--	--	--	--	---

VIEW ALL (1,117)

Now viewing project "My First Project" in organization "No organization" X of best-in-class

Subscribe

Google Cloud My First Project

Filter Type to filter

Category

- Analytics (1,117)
- Big data (584)
- Business Applications (209)
- Contact Center AI (13)
- Databases (504)

Analytics

 commercetools Reimaging digital commerce for the world's leading companies	 FullStory FullStory, INC Digital Experience Intelligence and Product Analytics	 LiveRamp LiveRamp Improves marketing by collaborating, joining, activating and measuring data across...	 Quantum Metric Quantum Metric The Platform For Continuous Product Design	 Shopify Shopify Inc The leading global commerce company that provides essential internet infrastructure for...	 eCommerce Search, Browse, & Recommendations Groupby B2B & B2C composable eCommerce Product Discovery Platform powered by Google Clou...
--	---	--	---	---	--

Type

- Google Cloud Platform (36)
- SaaS & APIs (1,648)
- Virtual machines (1,919)
- Data (165)
- Vertex AI (23)

Price

Analytics

 Apache Kafka® & Apache Flink® on Confluent Cloud™ Confluent Apache Kafka, reinvented for the modern data stack	 Databricks Databricks All your data, analytics and AI on one Lakehouse platform	 Fivetran Data Pipelines Fivetran Automated data integration for zero-maintenance data pipelines	 Informatica Intelligent Data Management Cloud Informatica Better Data. Even Smarter Analytics.	 Teradata VantageCloud Enterprise on Google Cloud Teradata The connected multi-cloud data platform for enterprise analytics.
---	--	--	---	---

DevOps

 Datadog Datadog See inside any stack, any app, at any scale, anywhere	 GitLab GitLab GitLab is a comprehensive AI-powered DevSecOps platform	 Dynatrace SaaS Dynatrace LLC Unified observability & security for cloud and on-premises	 Snyk: Developer Security Platform Snyk Ltd. Snyk: Developer Security Platform	 Harness Continuous Delivery Harness, Inc. Continuous Delivery as a Service
--	--	--	--	---

VIEW ALL (1,117)

Now viewing project "My First Project" in organization "No organization" X of best-in-class

≡ Google Cloud

← New Databricks subscription

Pricing Calculator

Order Summary

1. Select Plan

Plan
Databricks

Features

- Managed Apache Spark: Yes
- Optimized Delta Lake: Yes
- Cluster Autopilot: Yes
- Jobs Scheduling & Workflow: Yes
- Notebooks & Collaboration: Yes
- Databricks Runtime for ML: Yes
- Optimized Runtime Engine: Yes
- Administration Console: Yes
- Single Sign-On (SSO): Yes
- Role-based Access Control: Yes
- Token Management API: Yes

Pricing

Usage fee

Databricks Consumption Units <small>?</small>	USD 1.00
/unit	

Free
Estimated total cost

Adjust estimated timeframe

1 day 1 month 1 year

Monthly usage fee

Databricks Consumption Units

Estimated unit unit/mo USD 0.00/mo

2. Purchase details

Select a billing account * —
My Billing Account

3. Terms

Cancellation and change policy

- Your subscription fee is billed every month.

Additional terms

By purchasing, deploying, accessing, or using this product, you acknowledge that Google or an affiliate is the Vendor's agent with respect to this transaction and you agree to comply with the [Google Cloud Marketplace Terms of Service](#) (including the GPC terms set forth in Appendix A of the Marketplace ToS), [Databricks Terms of Service](#) and the terms of applicable open source software licenses bundled with the product.

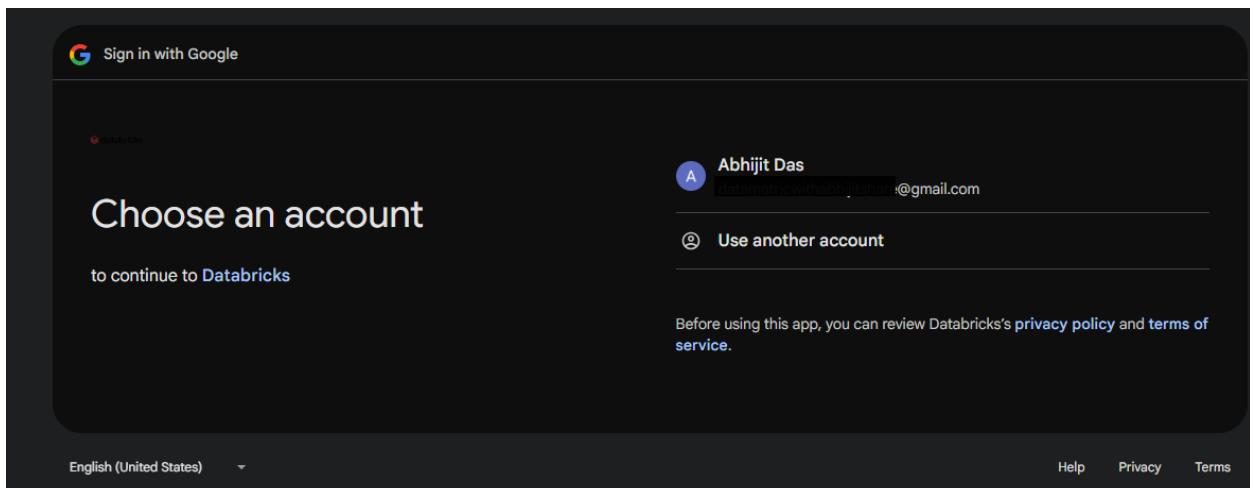
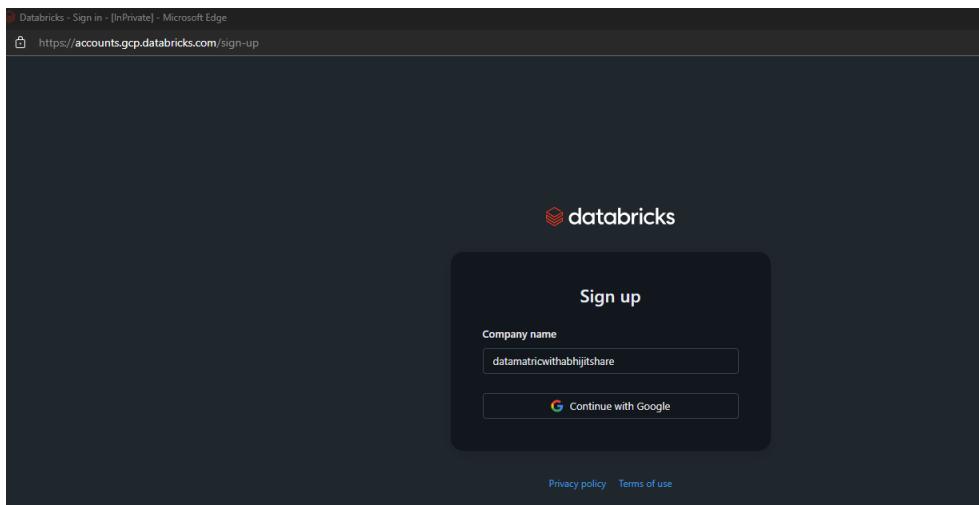
SUBSCRIBE

Your order request has been sent to Databricks



Your new subscription to the Databricks plan for Databricks requires your registration with Databricks. After Databricks approves your request, your subscription will be active and you will begin getting charged. This processing time will depend on the vendor, and you should reach out to Databricks directly with any questions related to signup.

MANAGE ORDERS **SIGN UP WITH DATABRICKS**



Click on Activate

A screenshot of the Google Cloud Billing interface. At the top, there is a message about a free trial status: "\$300.00 credit and 91 days remaining. Activate your full account to get unlimited access to all of Google Cloud—use any remaining credits, then pay only for what you use." On the right, there is a blue "Activate" button. The main area shows a list of orders under "Your orders for this product". One order is visible, with the provider listed as "Databricks". The status of this order is "Active". Other columns include Order number, Order title, Product, Plan, Next plan, Auto-renew, Purchase date, Start date, End date, and Payment schedule. At the bottom, there are buttons for "VIEW BILLING REPORTS", "VIEW BILLING COMMITMENTS", and "SEND FEEDBACK".

No need to activate the full account, unless you want to use any specific services.

Free trial status: \$300.00 credit and 88 days remaining. Activate your full account to get unlimited access to all of Google Cloud—use any remaining credits, then pay only for what you use.

Google Cloud

[Your orders for this product](#)

Select a billing account * — [My Billing Account](#)

This page contains all your orders for the Databricks product. To manage other product orders visit [Your orders](#).

Filter Filter by column name or chart value

Status	Order number	Order title	Provider	Product	Plan
<input checked="" type="checkbox"/> Active ⓘ	[REDACTED]	Databricks	Databricks	Databricks	[REDACTED]

Activate your full account

- Keep your cloud running uninterrupted
- Keep any remaining credits to spend during your Free Trial
- Pay only for what you use—billing starts once your Free Trial ends

[Cancel](#) [Activate](#)

Google Cloud [My First Project](#)

[Product details](#)

Databricks
Databricks

All your data, analytics and AI on one Lakehouse platform

Trial Active

[MANAGE ON PROVIDER](#) [CONTACT SALES](#) [Last purchased on 4/3/25](#)

[OVERVIEW](#) [PRICING](#) [DOCUMENTATION](#) [SUPPORT](#) [RELATED PRODUCTS](#)

Overview

Powered by Delta Lake, Databricks combines the best of data warehouses and data lakes into a lakehouse architecture, giving you one platform to collaborate on all of your data, analytics and AI workloads.

Notebooks: Build data science, data engineering and machine learning notebooks using Python, SQL, R, Scala. Collaborate on these notebooks with

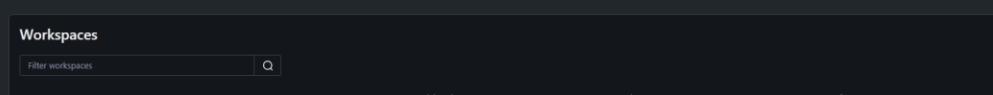
Additional details

Type: [SaaS & APIs](#)
 Last product update: 6/26/24
 Category: [Analytics](#), [Big data](#), [Machine learning](#)

The screenshot shows the Google Cloud Product details page for Databricks. At the top, there's a navigation bar with 'Google Cloud' and 'My First Project'. Below it, a back arrow and the text 'Product details'. The main content area features the Databricks logo and the tagline 'All your data, analytics and AI on one Lakehouse platform'. A green badge indicates 'Trial Active'. Below this, there are tabs for 'OVERVIEW', 'PRICING', 'DOCUMENTATION', 'SUPPORT', and 'RELATED PRODUCTS'. The 'OVERVIEW' tab is selected. A 'CONTACT SALES' button and a note 'Last purchased on 4/3/25' are visible. A modal window titled 'You're leaving Google' is open, stating 'This link opens a partner site. Google isn't responsible for privacy or security on third party sites.' It includes 'CANCEL' and 'OK' buttons. The background of the modal shows product details: Type: SaaS & APIs, Last product update: 6/26/24, Category: Analytics, Big data, Machine learning.

Click on “Continue With Google” account that we used for Databricks account registration and select the plan

The screenshot shows the 'Select a subscription plan' page from accounts.gcp.databricks.com. The URL in the address bar is https://accounts.gcp.databricks.com/subscription-plan-select?account_id=[REDACTED]. The page has a dark theme. At the top, it says 'Select a subscription plan'. Below that, a green button labeled 'Most popular' is highlighted. Underneath, there's a section for 'Premium' which includes 'Cloud native security and autoscaling' and three checked items: 'Databricks SQL', 'IP access list', and 'Role based access control'. A blue box highlights the 'Selected' status of the last item. Below this, a note says 'Your 14-day free trial starts when you click Continue. Thereafter, you will be charged at the list rates.' A large blue 'Continue' button is centered. At the bottom, there's a note about agreeing to Databricks terms and conditions, and a note about existing contracts.



The screenshot shows the Databricks Workspaces page. On the left, a sidebar menu includes 'Workspaces' (selected), 'Catalog', 'Usage', 'User management', 'Cloud resources', 'Previews', and 'Settings'. The main area is titled 'Workspaces' with a search bar and a 'Create workspace' button. A large central box contains a placeholder icon with arrows and the text 'Your workspace is the environment for doing work in Databricks. Create one to get started.' Below it is another 'Create workspace' button.

To create workspace, we need to fetch/copy GCP project-id.

The screenshot shows the GCP dashboard with a 'Welcome' message for user Abhiiit Das. A modal window titled 'Select a project' is open, listing two projects: 'GCP Databricks Project' and 'My First Project'. The 'RECENT' tab is selected. The background shows a sidebar with 'Recommendations' and 'Pre-built solutions' sections.

Name	ID
GCP Databricks Project	gcp-databricks-project-[REDACTED]
My First Project	modular-visitor-[REDACTED]

List all available regions and zones.

```
Enabling service [compute.googleapis.com] on project [gcp-databricks-project-15580]...
Operation "operations/acf.p2-[56300916766-0b1a15c9-dea7-4d1a-bea6-76f190237d90]" finished successfully.
NAME: africa-south1
CPUS: 0/24
DISKS_GB: 0/3277
ADDRESSES: 0/8
RESERVED_ADDRESSES: 0/8
STATUS: UP
TURNDOWN_DATE:

NAME: asia-east1
CPUS: 0/24
DISKS_GB: 0/3277
ADDRESSES: 0/8
RESERVED_ADDRESSES: 0/8
STATUS: UP
TURNDOWN_DATE:

NAME: asia-east2
CPUS: 0/24
DISKS_GB: 0/3277
ADDRESSES: 0/8
RESERVED_ADDRESSES: 0/8
STATUS: UP
TURNDOWN_DATE:
```

Command to find the region for a project in GCP.

```
gcloud compute project-info describe --project YOUR_PROJECT_ID
```

```
datamatricwthabhi@itsshare:~ (gcp-databricks-project-15580) $ gcloud compute project-info describe --project gcp-databricks-project-15580 | grep REGION
[ metric: CPUS_ALL_REGIONS
[ metric: GPUS_ALL_REGIONS ]
```

Enter the details to create Databricks – Workspace.

```
gcp_databricks_demo
```

The screenshot shows the 'Create Workspace' configuration page. On the left, a sidebar menu includes 'Workspaces' (selected), 'Catalog', 'Usage', 'User management', 'Cloud resources', 'Previews', and 'Settings'. The main form has 'Workspaces > Create workspace > Create Workspace' at the top. It asks for a 'Workspace name' (input: 'gcp_databricks_demo'), 'Region' (input: 'us-west1'), and 'Google cloud project ID' (input: 'gcp-databricks-project-155806'). Below these fields are two notes: one about creating Google Cloud Storage buckets and another about minimum resource quotas.

Create Workspace

Configurations

* **Workspace name**

`gcp_databricks_demo`

Human readable name for your workspace

* **Region**

`us-west1`

* **Google cloud project ID**

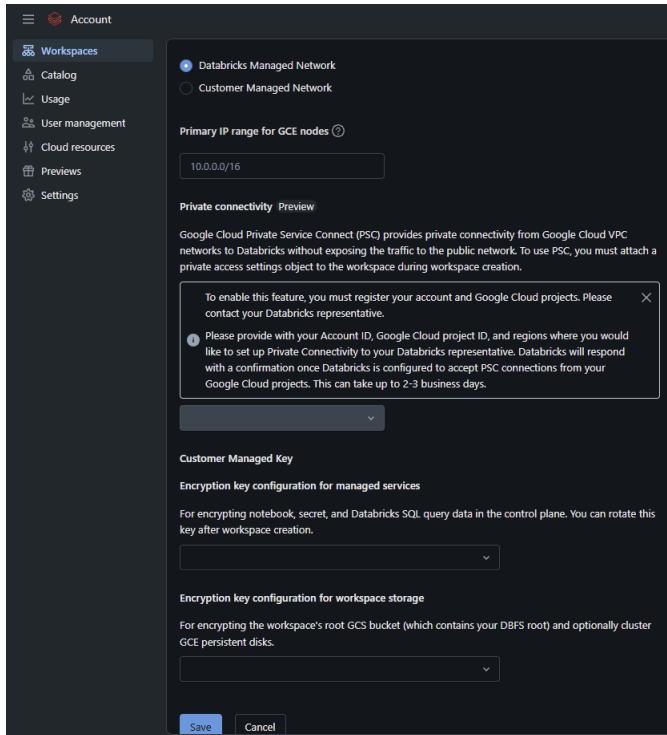
`gcp-databricks-project-155806`

Learn how to
find your Google Cloud project ID
or create a Google Cloud project if you do not
have one already.

Note
As part of creating this workspace, two Google Cloud Storage buckets will be created in your GCP project. These buckets will host the data that is put in the external DBFS storage and internal DBFS storage, respectively. Please review the access controls on these buckets in the GCP console. [Read more](#)

Note
please ensure that the project chosen has [the minimum resource quotas](#) to run your jobs and clusters.

We will keep other details unchanged for the demo.

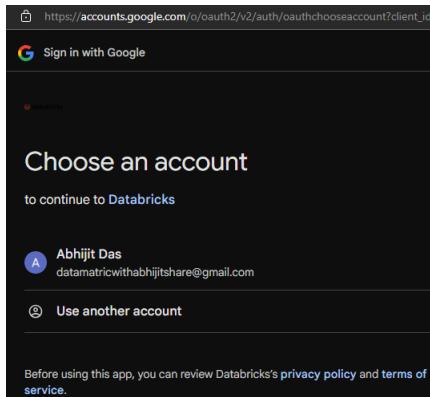


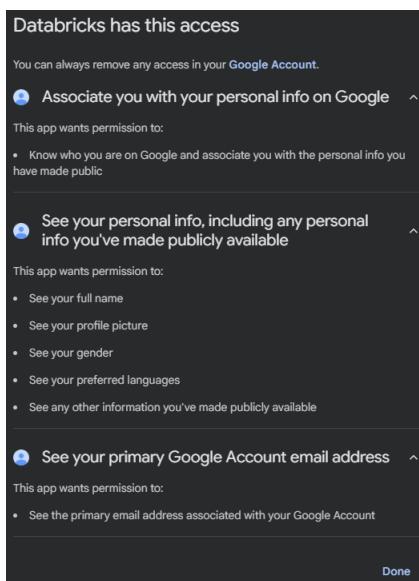
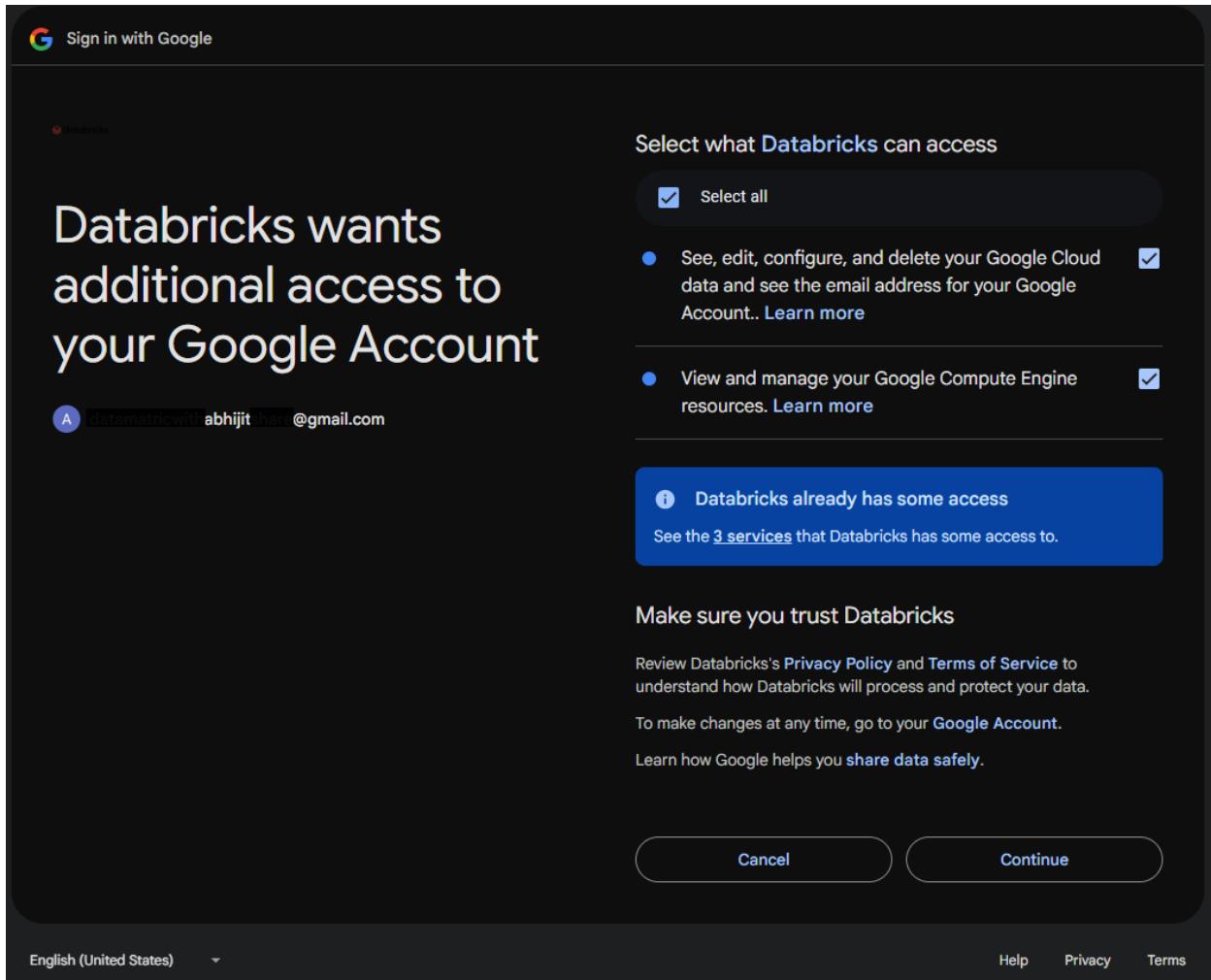
Click on “Save”

Note: It is important to understand quotas to run jobs and clusters with GCP Databricks. To know more, you can click on this [link](#).

Two GCS buckets will be automatically created in your GCP project. These buckets will host the data that is put in the external DBFS storage and internal DBFS storage, respectively. To understand access controls on these buckets in the GCP console, you can look into this [link](#).

It will open a window to authenticate your credentials. Select appropriate gmail-account.





Now your Databricks workspace is getting provisioned

The screenshot shows the Databricks account interface with the 'Workspaces' section selected. A table lists one workspace:

Name	Status	Pricing tier	Region	Created
gcp_databricks_demo	Provisioning	Premium	us-west1	today at 10:02 PM

Once it is in running state.

The screenshot shows the Databricks account interface with the 'Workspaces' section selected. The workspace status has changed to 'Running':

Name	Status	Pricing tier	Region	Created
gcp_databricks_demo	Running	Premium	us-west1	today at 10:02 PM

Unity-Catalog will be created as well.

The screenshot shows the Databricks account interface with the 'Catalog' section selected. A table lists one metastore:

Name	Region	Path	Created at	Updated at
metastore_gcp_us_west1	us-west1		today at 10:04 PM	today at 10:04 PM

Usage page (verify usage and total USD/DBU spent).

The screenshot shows the Databricks account interface with the 'Usage' section selected. The 'Consumption (Legacy)' tab is active. A message indicates 'No usage data found in the given time range'. Below this, there is a 'Usage details' section with a table:

Name	Total dollars spent	Jobs	All purpose	SQL compute	Serverless SQL compute	Delta live tables

User management (Users, SPs and Groups)

User management

Users **Service principals** **Groups**

Add users to the account. Account users can use the account console to view and connect to their workspaces. Account admins can perform all of the management functions available in the account console. [Learn more](#)

Status	Name	Email	Roles
Green	Abhijit Das	abhijit@gmail.com	Account admin

For Network setup (config, private access and VPC endpoints) and Encryption Key Config

Cloud resources

Network [Encryption keys configuration](#)

To enable Databricks Private Service Connect, you must register your account and Google Cloud projects. Please contact your Databricks representative.

Please provide with your Account ID, Google cloud project ID, and regions where you would like to set up Private Connectivity to your Databricks representative. Databricks will respond with a confirmation once Databricks is configured to accept PSC connections from your Google Cloud projects. This can take up to 2-3 business days.

Network configurations

Private access settings

VPC endpoints

Name	VPC ID	Created

To know your current plan, user provisioning, tokens reports, app connections and other account related activities, you need to look into Settings

Settings

[Subscription & billing](#) [Log delivery](#) [User provisioning](#) [Token report](#) [App connections](#) [Feature enablement](#) [Language settings](#) [My preferences](#) [Security and compliance](#) [Account settings](#)

Subscription

Premium

Billing information

Your subscription will be billed through Google Cloud Marketplace.

[Change plan](#)

Billing

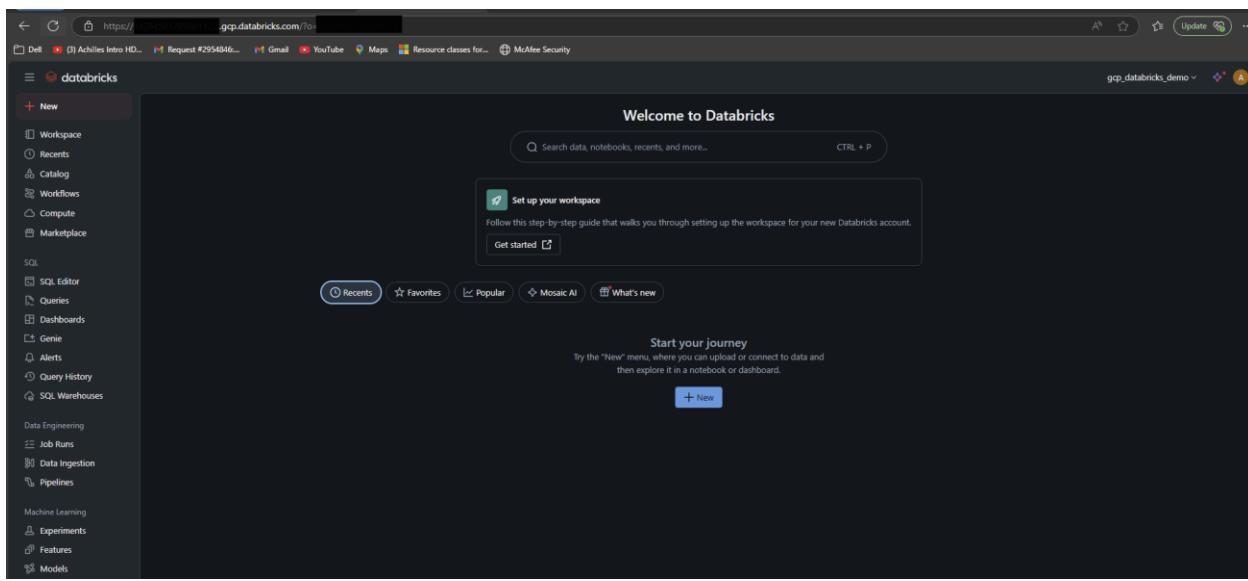
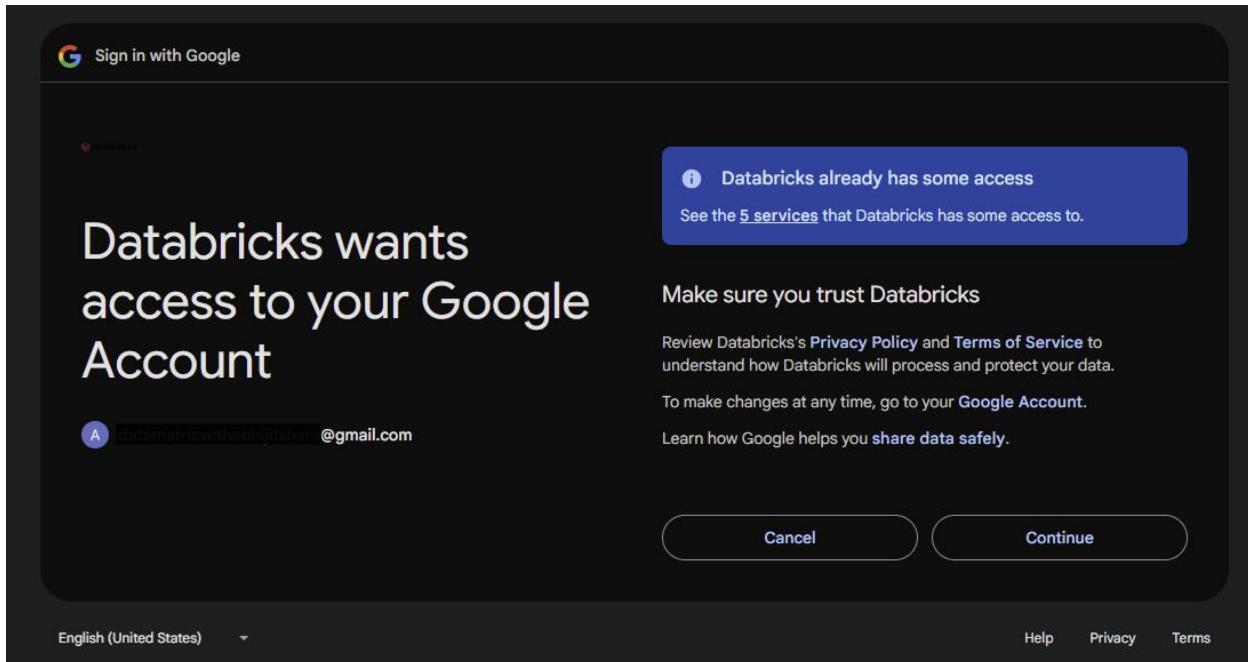
Order ID

Order ID can be used to associate your Google Billing account with your Databricks account. [Learn more](#)

Open workspace

The screenshot shows the Databricks Configuration page for the workspace 'gcp_databricks_demo'. On the left, there's a sidebar with options like Workspaces, Catalog, Usage, User management, Cloud resources, Previews, and Settings. The main area displays configuration details for GCP storage buckets. It shows a table with three columns: Google Cloud project ID (containing 'gcp-databricks-project-'), Region ('us-west1'), and Metastore ('metastore_gcp_us_west1'). A note at the top states: 'Note: As part of this workspace, two Google Cloud Storage buckets have been created in your GCP project. These buckets host the data that is put in the external DBFS storage and internal DBFS storage, respectively. Please review the access controls on these buckets in the GCP console.' Below this, there's a section for workspace status showing 'Running' and 'Workspace is running.' At the bottom right, there's a 'Open workspace' button.

The screenshot shows the 'Choose an account' screen for Databricks. It features a dark background with white text. At the top left is a 'Sign in with Google' button. The main heading 'Choose an account' is centered above the instruction 'to continue to Databricks'. To the right, there's a profile card for 'Abhijit Das' with a blue circular icon containing a white letter 'A', the name 'Abhijit Das', and the email 'abhinavdas1991@gmail.com'. Below this is a link 'Use another account' with a user icon. At the bottom, a note reads: 'Before using this app, you can review Databricks's [privacy policy](#) and [terms of service](#)'. The footer includes language selection ('English (United States)'), and links for 'Help', 'Privacy', and 'Terms'.



Now that Databricks workspace has been created. We can go ahead and create cluster. Next we will connect to Google Cloud Storage.

To create cluster, we need to navigate to Compute and “Create Compute”,

Note: if you are facing issue with quota (common issue), then you need go to the google-cloud account -> quota and system limit -> search for CPUs - "CPUs for N2", "CPUs for E2", CPUs for C2, or whatever VM family your cluster is using (you can check the node type in the cluster settings)

I was unable to find E2 family cluster and it was throwing error from Databricks Compute

"This account may not have enough CPU cores to start a cluster. Contact your administrator to increase the limits."

I was able to see A2 and C2 family CPUs with Value – 8. If you are looking for any specific family, then you can request for a quota increase.

Connect to Google Cloud Storage: <https://docs.databricks.com/gcp/en/connect/storage/gcs>

Set up Google Cloud service account using Google Cloud Console

The screenshot shows the Google Cloud IAM & Admin / Service accounts page. The left sidebar is titled "Google Cloud" and includes sections for IAM, PAM, Principal Access Boundary, Organizations, Identity & Organization, Policy Troubleshooter, Policy Analyzer, Organization Policies, and Service Accounts (which is currently selected). The main content area has a header "Service accounts" with buttons for "+ Create service account", "Delete", "Manage access", and "Refresh". Below this, a section titled "Service accounts for project 'GCP Databricks Project'" provides information about service accounts and organization policies. A table lists two service accounts:

Email	Status	Name	Description	Key ID	Key creation date	OAuth 2 Client ID	Actions
compute@developer.gserviceaccount.com	Enabled	Compute Engine default service account	No keys				⋮
databricks-compute@gcp-databricks-project-iam.gserviceaccount.com	Enabled	Databricks Compute Service Account	Default service account attached to VMs for Databricks clusters with no custom service account				⋮

[!\[\]\(4cb8b8eab030c2bacbee634b62ce982f_img.jpg\) Create service account](#)

1 Service account details

Service account name

databricks-gcs-mount

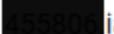
Display name for this service account

Service account ID *

databricks-gcs-mount

Email address:  @gcp-databricks-project-

 iam.gserviceaccount.com 

Service account description

Setup Databricks GCS Connection For Mount

Describe what this service account will do

[Create and continue](#)

**2 Grant this service account access to project
(optional)**

**3 Grant users access to this service account
(optional)**

[Done](#)

[Cancel](#)

Optional

Create service account

Service account details

Grant this service account access to project (optional)

Grant this service account access to GCP Databricks Project so that it has permission to complete specific actions on the resources in your project.

[Learn more](#)

Select a role

Filter Filter by role or permission

Quick access

- Currently used
- Custom
- Basic

By product or service

- Access Approval
- Access Context

Done

Service account details

Grant this service account access to project (optional)

Grant users access to this service account (optional)

Grant access to users or groups that need to perform actions as this service account.

[Learn more](#)

Service account users role

Grant users the permissions to deploy jobs and VMs with this service account

Service account admins role

Grant users the permission to administer this service account

Done **Cancel**

Service Account has been created. Click on this service account to get the email-id for Databricks Compute Cluster

Service accounts								+ Create service account	Delete	Manage access	Refresh	Learn
Service accounts for project "GCP Databricks Project"												
A service account represents a Google Cloud service identity, such as code running on Compute Engine VMs, App Engine apps, or systems running outside Google. Learn more about service accounts												
Organization policies can be used to secure service accounts and block risky service account features, such as automatic IAM Grants, key creation/upload, or the creation of service accounts entirely. Learn more about service account organization policies												
<input type="checkbox"/> Email		Status	Name	Description	Key ID	Key creation date	OAuth 2 Client ID					
<input type="checkbox"/> compute@developer.gserviceaccount.com			Compute Engine default service account		No keys							
<input type="checkbox"/> databricks-compute@gcp-databricks-project-[REDACTED].iam.gserviceaccount.com			Databricks Compute Service Account	Default service account attached to VMs for Databricks clusters with no custom service account	No keys							
<input type="checkbox"/> databricks-gcs-mount@gcp-databricks-project-[REDACTED].iam.gserviceaccount.com			databricks-gcs-mount	Setup Databricks GCS Connection For Mount	No keys							

The screenshot shows the Google Cloud IAM & Admin interface. On the left, a sidebar lists various administrative sections: IAM, PAM, Principal Access Boundary, Organizations, Identity & Organization, Policy Troubleshooter, Policy Analyzer, Organization Policies, Service Accounts (which is selected and highlighted in blue), Workload Identity Federat..., Workforce Identity Federa..., Labels, Tags, Settings, Privacy & Security, Identity-Aware Proxy, Roles, Manage Resources, and Release Notes. The main content area is titled 'databricks-gcs-mount' and contains tabs for Details, Permissions, Keys, Metrics, and Logs. The Details tab is active. It shows the service account's name as 'databricks-gcs-mount', its description as 'Setup Databricks GCS Connection For Mount', its email as 'databricks-gcs-mount@gcp-databricks-project-[REDACTED].iam.gserviceaccount.com', and its unique ID as '[REDACTED]'. Below this, the 'Service account status' section indicates it is enabled and provides a 'Disable service account' button. The 'Advanced settings' section includes a warning about domain-wide delegation, stating: 'Granting this service account access to your organization's data via domain-wide delegation should be used with caution. It can be reversed by disabling or deleting the service account or by removing access through the Google Workspace admin console.' A link 'Learn more about domain-wide delegation' is provided.

IAM & Admin / Service accounts / Service account: [REDACTED]

databricks-gcs-mount

Details Permissions Keys Metrics Logs

Service account details

Name: databricks-gcs-mount Save

Description: Setup Databricks GCS Connection For Mount Save

Email: databricks-gcs-mount@gcp-databricks-project-[REDACTED].iam.gserviceaccount.com

Unique ID: [REDACTED]

Service account status

Disabling your account allows you to preserve your policies without having to delete it.

Enabled

[Disable service account](#)

Advanced settings

Domain-wide Delegation

⚠️ Granting this service account access to your organization's data via domain-wide delegation should be used with caution. It can be reversed by disabling or deleting the service account or by removing access through the Google Workspace admin console.

[Learn more about domain-wide delegation](#)

Next Step - Configure your GCS bucket

Note: To work with DBFS mounts, your bucket name must not contain an underscore.

Google Cloud GCP Databricks Project Search (/) for resources, docs, products, and more Search

Cloud Storage Overview

Welcome back to Cloud Storage, Abhijit!

Pinned buckets

You haven't pinned any buckets yet. To pin a bucket for quick access, click the pin icon next to the bucket's name within the bucket, or select the pin option from the bucket's 3-dot overflow menu on the bucket list.

[Go to all buckets](#)

Recently visited

You haven't visited any buckets or folders recently. Once you have viewed any Cloud Storage buckets or folders, a list of history will be displayed for easy access in the future.

Recent transfer activity

There's no recent transfer activity. Use Storage Transfer Service jobs to move data into or between your Cloud Storage buckets. [Learn more](#)

[Create a transfer job](#)

Google Cloud GCP Databricks Project Search (/) for resources, docs, products, and more Search

Buckets

Create Refresh

Name	Created	Location type	Location	Default storage class	Last modified	Public access	Access
databricks[REDACTED]	Apr 6, 2025, 10:03:49 PM	Region	us-west1	Standard	Apr 6, 2025, 10:03:49 PM	Not public	Unif
databricks[REDACTED]system	Apr 6, 2025, 10:03:48 PM	Region	us-west1	Standard	Apr 6, 2025, 10:03:48 PM	Not public	Unif
databricks[REDACTED]unity	Apr 6, 2025, 10:03:50 PM	Region	us-west1	Standard	Apr 6, 2025, 10:03:50 PM	Not public	Unif

Google Cloud GCP Databricks Project Search (/) for resources, docs, products, and more Search

Create a bucket

Get Started

Name: sourcedriverdatademo

Choose where to store your data

Location: us (multiple regions in United States)
Location type: Multi-region

Choose a storage class for your data

Default storage class: Standard

Choose how to control access to objects

Public access prevention: Off
Access control: Uniform

Choose how to protect object data

Your data is always protected with Cloud Storage but you can also choose from these additional data protection options to add extra layers of security.

Data protection

Soft delete policy (For data recovery)
When enabled, this bucket and its objects will be kept for a specified period after they're deleted and can be restored during this time. [Learn more](#)

Object versioning (For version control)
For restoring deleted or overwritten objects. To minimize the cost of storing versions, we recommend limiting the number of noncurrent versions per object and scheduling them to expire after a number of days. [Learn more](#)

Retention (For compliance)
For preventing the deletion or modification of the bucket's objects for a specified period of time.

Data encryption

Google-managed encryption key
Keys owned by Google

Cloud KMS key

Good to know

Location pricing
Storage rates vary depending on the storage class of your data and location of your bucket. [Pricing details](#)

Current configuration: Multi-region / Standard

Item	Cost
us (multiple regions in United States)	\$0.026 per GB-month
With default replication	\$0.020 per GB written

Estimate your monthly cost

Configure the bucket:

The screenshot shows the Google Cloud Storage Bucket Details page for a bucket named "sourcedriverdatademo". The left sidebar includes links for Overview, Buckets (which is selected), Monitoring, Settings, Storage Intelligence, Insights datasets, and Configuration. The main content area displays bucket metadata: Location (us (multiple regions in United States)), Storage class (Standard), Public access (Not public), and Protection (None). Below this, tabs for Objects, Configuration, Permissions (selected), Protection, Lifecycle, Observability (New), Inventory Reports, and Operations are shown. The Permissions section contains two main boxes: "Public access" (Not public) and "Access control" (Uniform: No object-level ACLs enabled). It also includes a note about 90 days left to change the setting and a link to switch to fine-grained access. At the bottom, there are "View by principals" and "View by roles" buttons, along with "Grant access" and "Remove access" links.

Provide the desired permission to the service account on the bucket from the Cloud Storage roles - > Storage Admin: Grants full privileges on this bucket.

Grant access to "sourcedriverdatademo"

Grant principals access to this resource and add roles to specify what actions the principals can take. Optionally, add conditions to grant access to principals only when a specific criteria is met. [Learn more about IAM conditions](#)

Resource

 sourcedriverdatademo

Add principals

Principals are users, groups, domains, or service accounts. [Learn more about principals in IAM](#)

New principals *

X?

Assign roles

Roles are composed of sets of permissions and determine what the principal can do with this resource. [Learn more](#)

Role * ▼ IAM condition (optional) ? + Add IAM condition trash bin

Grants full control of buckets and objects.

[+ Add another role](#)

[Save](#)

[Cancel](#)

Optional: Add keys for global configuration (Click Keys. The Google Cloud console displays a list of keys for the service account, including metadata for each key)

Note: We can mount GCS bucket into Databricks without using private-keys.

Private key saved to your computer

⚠ gcp-databricks-project-[REDACTED].json allows access to your cloud resources, so store it securely. [Learn more best practices](#)

[close](#)

JSON file will be downloaded.

Type	Status	Key	Creation date	Expiration date	
<input checked="" type="radio"/>	✓ Active	[REDACTED]	Apr 11, 2025	Dec 31, 9999	trash can

Configure a Databricks cluster -> When you configure your cluster, expand Advanced and set the Google Service Account field to your service account email address.

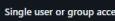
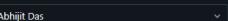
Compute > New compute > Simple form: OFF ▾

GCP Databricks Cluster

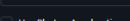
⚠ Based on your existing cloud quotas, your default compute size has been reduced. View or increase your quotas on GCP Cloud Console. [View](#) 

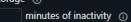
UI | [JSON](#)

Policy  
 Multi node Single node

Access mode  
 Dedicated (formerly: Single user)   Abhijit Das

Performance

Databricks runtime version  
 Use Photon Acceleration 

Node type  
 Enable autoscaling local storage 
 Terminate after  minutes of inactivity 

Summary

1 Driver	16 GB Memory, 4 Cores
Runtime	15.4.x-scala2.12
Unity Catalog	c3-standard-4-lssd 0.85 DBU/h

Compute > New compute > Simple form: OFF

GCP Databricks Cluster

Use Photon Acceleration 

Node type

c3-standard-4-lssd

16 GB Memory, 4 Cores  

Enable autoscaling local storage 

Terminate after minutes of inactivity 

Tags

Add tags

Key

Value

> Automatically added tags

▼ Advanced options

Google Service Account

databricks-gcs-mount@gcp-databricks-project-455806.iam.gserviceaccount.com

Instances

Spark

Logging

Init Scripts

Availability zone

auto

local SSDs

Default

Each local SSD is 375GB in size

Create compute

Cancel

GCP Databricks Cluster

Configuration

- Policy: Unrestricted
- Access mode: Single user or group access (Dedicated)
- DataBricks Runtime Version: 15.4 LTS (includes Apache Spark 3.5.0, Scala 2.12)
- Node type: c3-standard-4-ssd (16 GB Memory, 4 Cores)
- Tags: No custom tags

After starting the cluster, we should be able to mount GCS Bucket and read files.

GCSMountTest Python

File Edit View Run Help Last edit was 2 minutes ago

Run all GCP Databricks Cluster Schedule Share

2

```
bucket_name = "sourcedriverdatademo"
mount_name = "gcmount"
dbutils.fs.mount(
    f"gs://{bucket_name}",
    f"/mnt/databricks/{mount_name}",
    extra_configs = {"fs.gs.project.id": "gcp-databricks-project-[REDACTED]"}
)
```

True

3

```
%fs
ls /mnt/databricks/gcmount/csv_files/
```

path	name	size	modificationTime
dbfs:/mnt/databricks/gcmount/csv_files/circuits.csv	circuits.csv	10044	1744357388062

1 row | 0.47s runtime Refreshed 9 minutes ago

4

```
df_csv = spark.read.format("csv").option("header", "true").load("/mnt/databricks/gcmount/csv_files/circuits.csv")
display(df_csv)
```

(2) Spark Jobs

df_csv: pyspark.sql.dataframe.DataFrame = [circuitId: string, circuitRef: string ... 7 more fields]

circuitId	circuitRef	name	location	country	lat	long	alt	url
-----------	------------	------	----------	---------	-----	------	-----	-----

2 minutes ago (2s) 4

```
df_csv = spark.read.format("csv").option("header", "true").load("/mnt/databricks/gcsmount/csv_files/circuits.csv")
display(df_csv)
```

(2) Spark Jobs

df_csv: pyspark.sql.dataframe.DataFrame = [circuitid: string, circuitRef: string ... 7 more fields]

Table +

	circuitid	circuitRef	name	location	country	lat	lng	alt	url
6	6	monaco	Circuit de Monaco	Monte-Carlo	Monaco	43.7347	742056	7	http://en.wikipedia.org/wiki/Circuit_de_Monaco
7	7	villeneuve	Circuit Gilles Villeneuve	Montreal	Canada	45.5	-73.5228	13	http://en.wikipedia.org/wiki/Circuit_Gilles_Villeneuve
8	8	magny_cours	Circuit de Nevers Magny-Cours	Magny Cours	France	46.8642	3.16361	228	http://en.wikipedia.org/wiki/Circuit_de_Nevers_Magny_Cours
9	9	silverstone	Silverstone Circuit	Silverstone	UK	52.0786	-1.01694	153	http://en.wikipedia.org/wiki/Silverstone_Circuit
10	10	hockenheimring	Hockenheimring	Hockenheim	Germany	49.3278	8.56583	103	http://en.wikipedia.org/wiki/Hockenheimring
11	11	hungaroring	Hungaroring	Budapest	Hungary	47.5789	19.2486	264	http://en.wikipedia.org/wiki/Hungaroring
12	12	valencia	Valencia Street Circuit	Valencia	Spain	39.4589	-0.331667	4	http://en.wikipedia.org/wiki/Valencia_Street_Circuit
13	13	spa	Circuit de Spa-Francorchamps	Spa	Belgium	50.4372	5.97139	401	http://en.wikipedia.org/wiki/Circuit_de_Spa-Francorchamps
14	14	monza	Autodromo Nazionale di Monza	Monza	Italy	45.6156	9.28111	162	http://en.wikipedia.org/wiki/Autodromo_Nazionale_di_Monza
15	15	marina_bay	Marina Bay Street Circuit	Marina Bay	Singapore	1.2914	103.864	18	http://en.wikipedia.org/wiki/Marina_Bay_Street_Circuit
16	16	fuji	Fuji Speedway	Oyama	Japan	35.3717	138.927	583	http://en.wikipedia.org/wiki/Fuji_Speedway
17	17	shanghai	Shanghai International Circuit	Shanghai	China	31.3389	121.22	5	http://en.wikipedia.org/wiki/Shanghai_International_Circuit
18	18	interlagos	Autódromo José Carlos Pace	São Paulo	Brazil	-23.7036	-46.6997	785	http://en.wikipedia.org/wiki/Aut%C3%B3dromo_Jos%C3%A9_Carlos_Pace
19	19	indyapolis	Indianapolis Motor Speedway	Indianapolis	USA	39.795	-86.2347	223	http://en.wikipedia.org/wiki/Indianapolis_Motor_Speedway
20									

77 rows | 1.96s runtime

Refreshed 2 minutes ago

Creating a Delta Live Table with Pipeline in GCS.

Delta Live Pipeline Python

File Edit View Run Help Last edit was 2 minutes ago

Step 1: Bronze Layer - Declare the source table from GCS-mounted CSV

```
2 minutes ago (2s) 2
```

```
import dlt
from pyspark.sql.functions import *

@dlt.table(
    name="bronze_circuits",
    comment="Raw circuits data from GCS",
    table_properties={"quality": "bronze"}
)
def load_raw_circuits_data():
    return (
        spark.read
            .option("header", True)
            .csv("/mnt/databricks/gcsmount/csv_files/circuits.csv")
    )
```

(1) Spark Jobs

bronze_circuits is defined as a Delta Live Tables dataset with schema:

Name	Type
circuitid	string
circuitRef	string
name	string
location	string
country	string
lat	string
lng	string
alt	string
url	string

To populate your table you must either:

- Run an existing pipeline using the Delta Live Tables menu
- Create a new pipeline [Create Pipeline](#)

Creating a DLT Pipeline with Unity Catalog (in preview now). As it is a batch job, so we will be using triggered mode instead of Continuous.

Workflows > Pipelines >

Create pipeline [Send feedback](#)

[UI](#) [JSON](#)

General

* Pipeline name

Product edition [Help me choose](#)

Pipeline mode Triggered Continuous

Source code

Paths [Edit](#)

Add source code

If you don't add any source code, Databricks will create an empty notebook for the pipeline. You can edit this notebook later.

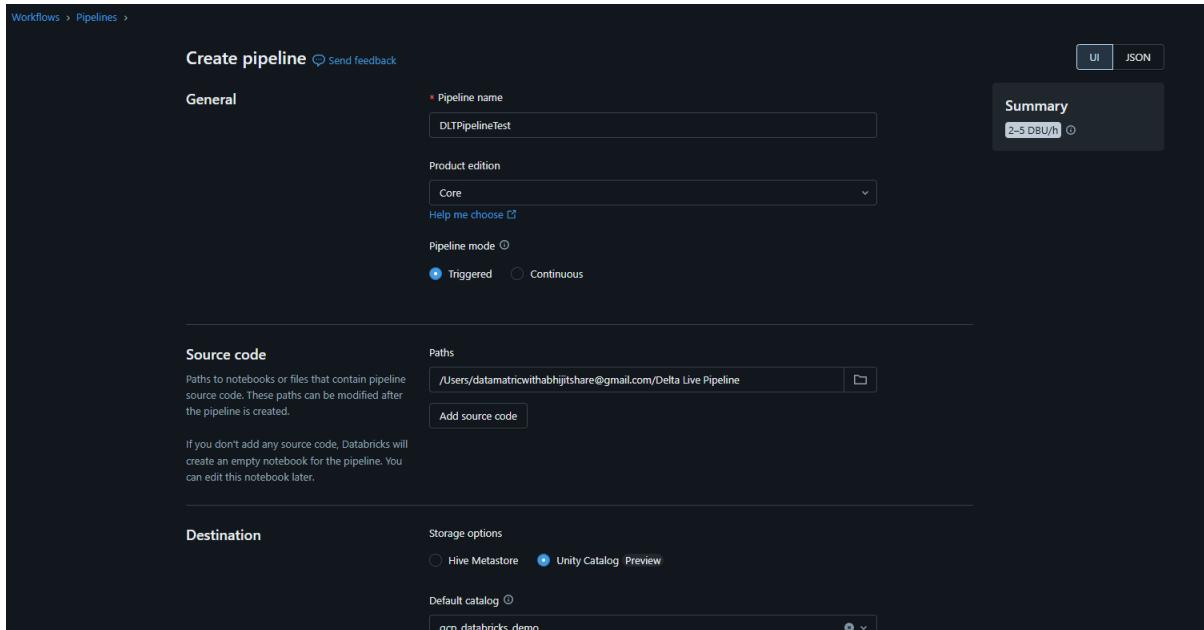
Destination

Storage options Hive Metastore Unity Catalog [Preview](#)

Default catalog [Edit](#)

Summary

2-5 DBU/h [Edit](#)



Workflows > Pipelines >

Create pipeline [Send feedback](#)

[UI](#) [JSON](#)

Destination

Storage options Hive Metastore Unity Catalog [Preview](#)

Default catalog [Edit](#)

* Default schema [Edit](#)

Compute

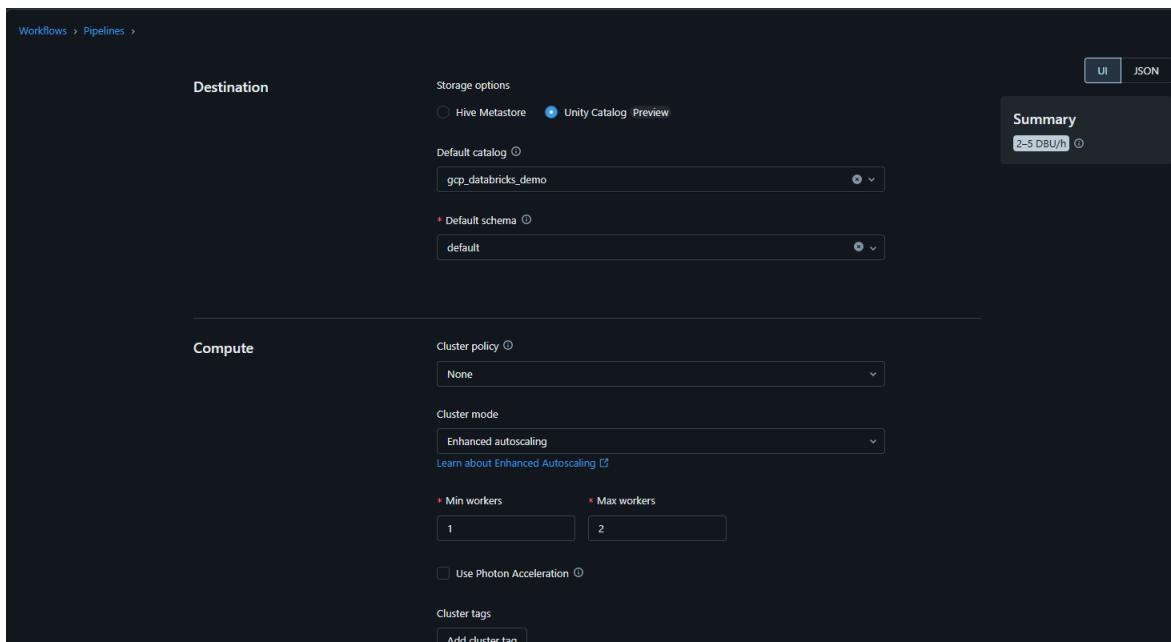
Cluster policy None [Edit](#)

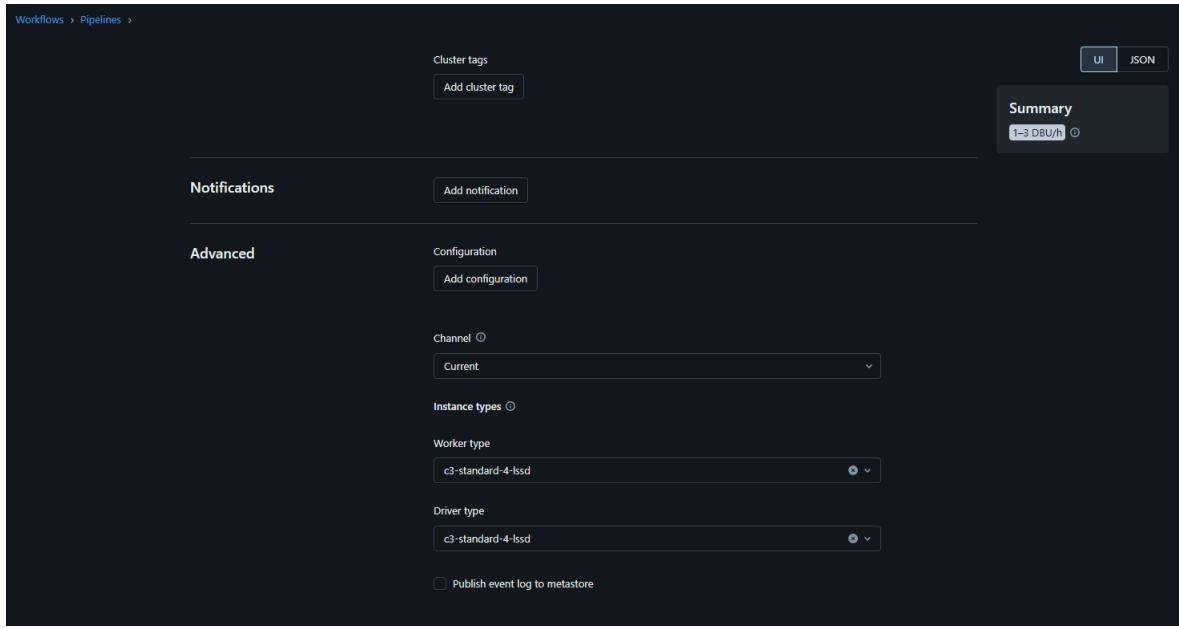
Cluster mode Enhanced autoscaling [Learn about Enhanced Autoscaling](#)

* Min workers * Max workers

Use Photon Acceleration [Edit](#)

Cluster tags [Edit](#)





I was able to reduce number of DBU use by selecting better worker and driver types. As I will be using this pipeline for a quick demo, so I tried to keep low configuration.

The screenshot shows the 'Delta Live Pipeline' editor with a single step defined:

Step 1: Bronze Layer - Declare the source table from GCS-mounted CSV

```
import dlt
from pyspark.sql.functions import *

@dlt.table(
    name="bronze_circuits",
    comment="Raw circuits data from GCS",
    table_properties={"quality": "bronze"}
)
def load_raw_circuits_data():
    return (
        spark.read
        .option("header", True)
        .csv("/mnt/databricks/gcsmount/csv_files/circuits.csv")
    )
```

Below the code, there is a note about keyboard shortcuts:

[Shift+Enter] to run and move to next cell
[Ctrl+Shift+P] to open the command palette
[Esc H] to see all keyboard shortcuts

At the bottom, it says: 'A graph will be generated here once a pipeline update has started. Click "Start" to start an update.'

Validating the pipeline

```

import dlt
from pyspark.sql.functions import *

@dlt.table(
    name="bronze_circuits",
    comment="Raw circuits data from GCS",
    table_properties={"quality": "bronze"}
)
def load_raw_circuits_data():
    return (
        spark.read
        .option("header", True)
        .csv("/mnt/databricks/gcsmount/csv_files/circuits.csv")
    )

```

[Shift+Enter] to run and move to next cell
[Ctrl+Shift+P] to open the command palette
[Esc H] to see all keyboard shortcuts

DLT graph DLT event log DLT query history

```

graph TD
    1[Creating validate-only update] --> 2[Waiting for resources]
    2 --> 3[Initializing]
    3 --> 4[Rendering graph]

```

Integrate Databricks With BigQuery

Connect to BigQuery Storage API - Enable the [BigQuery Storage API](#).

BigQuery Storage API
Google Enterprise APIs

MANAGE API Enabled

OVERVIEW PRICING DOCUMENTATION RELATED PRODUCTS

Overview

Additional details

- Type: [SaaS & APIs](#)
- Last product update: 7/21/22
- Category: [Google Enterprise APIs](#)
- Service name: [bigquerystorage.googleapis.com](#)

Pricing

BigQuery Storage API Network Data Transfer Out via Carr	USD 0.042
BigQuery Storage API Network Data Transfer Out via Carr	/gibibyte
BigQuery Storage API Network Data Transfer Out via Carr	
BigQuery Storage API Network Inter Region Data Transfer	

Note: Additional charges may also apply from infrastructure you use to call the API. If you pay in a currency other than USD, the prices listed in your currency on [Google Cloud](#)

Create a service account to integrate Databricks cluster with BigQuery.

Google Cloud GCP Databricks Project servi

IAM & Admin / Service accounts / Create service account

Create service account

1 Service account details

Service account name: databricks-bigquery-demo

Display name for this service account:

Service account ID *: databricks-bigquery-demo@gcp-databricks-project-455806.iam.gserviceaccount.com

Email address: databricks-bigquery-demo@gcp-databricks-project-455806.iam.gserviceaccount.com

Service account description: Databricks BigQuery service account

Describe what this service account will do:

Create and continue

2 Grant this service account access to project (optional)

3 Grant users access to this service account (optional)

Done Cancel

To give permission to write data, grant the following roles:

Google Cloud GCP Databricks Project servi

IAM & Admin / Service accounts / Create service account

Create service account

1 Service account details

Grant this service account access to project (optional)

Grant this service account access to GCP Databricks Project so that it has permission to complete specific actions on the resources in your project. Learn more

Role: BigQuery Data Editor IAM condition (optional) + Add IAM condition

Access to edit all the contents of datasets

Role: BigQuery Job User IAM condition (optional) + Add IAM condition

Access to run jobs

+ Add another role

Continue

3 Grant users access to this service account (optional)

Done Cancel

Google Cloud GCP Databricks Project Search (/) for resources, docs, products, and more

Cloud Storage Create a bucket

Overview Buckets Monitoring Settings

Storage Intelligence Insights datasets Configuration

Get Started Pick a globally unique, permanent name. [Naming guidelines](#)

databricks-bq-demo

Tip: Don't include any sensitive information

Optimize storage for data-intensive workloads

Labels (optional)

Continue

Good to know

Location pricing Storage rates vary depending on the storage class of your data and location of your bucket. [Pricing details](#)

Current configuration: Multi-region / Standard

Item	Cost
us (multiple regions in United States)	\$0.026 per GB-month
With default replication	\$0.020 per GB written

Estimate your monthly cost

Choose where to store your data

Location: us (multiple regions in United States)
Location type: Multi-region

Choose a storage class for your data

Default storage class: Standard

Choose how to control access to objects

Public access prevention: On
Access control: Uniform

Choose how to protect object data

Soft delete policy: Default
Object versioning: Disabled
Bucket retention policy: Disabled
Object retention: Disabled
Encryption type: Google-managed

Marketplace Release Notes

Create Cancel

Cluster update

Compute > Simple form: OFF ▾

Abhijit Das's Cluster • GCE

[Configuration](#) [Notebooks \(1\)](#) [Libraries](#) [Event log](#) [Spark UI](#) [Driver logs](#) [Metrics](#) [Apps](#) [Spark compute UI - Master](#)

Dedicated (formerly, single user) [Abhijit Das](#)

Performance

Databricks Runtime Version

15.4 LTS (includes Apache Spark 3.5.0, Scala 2.12)

Use Photon Acceleration ⓘ

Node type ⓘ

c3-standard-4-1ssd 16 GB Memory, 4 Cores

Enable autoscaling local storage ⓘ

Terminate after minutes of inactivity ⓘ

Tags ⓘ

No custom tags

> Automatically added tags

▼ Advanced options

Google Service Account ⓘ

databricks-bigquery-demo@gcp-databricks-project-455806.iam.gserviceaccount.com

[Instances](#) [Spark](#) [Logging](#) [Init Scripts](#) [JDBC/ODBC](#)

Availability zone

auto

local SSDs

Default Each local SSD is 375GB in size

After updating the Databricks Cluster, we tried to read a sample data. We were getting error below.

Untitled Notebook 2025-04-17 23:52:59 Python ▾ [Share](#)

File Edit View Run Help Last edit was 1 minute ago

Run all Abhijit Das's Cluster Schedule

```

1 1 minute ago (5s)
  table = "Bigquery-public-data.samples.shakespeare"
  df = spark.read.format("bigquery").option("table",table).load()
  df.createOrReplaceTempView("shakespeare")
  df: pyspark.sql.DataFrame = [word: string, word_count: long ... 2 more fields]

2 Last execution failed
  1 %sql
  2 select * from shakespeare
  ① > PermissionDeniedException: com.google.cloud.spark.bigquery.repackaged.io.grpc.StatusRuntimeException: PERMISSION_DENIED: request failed: the user does not have 'bigquery.readsessions.create' permission for 'projects/gcp-databricks-project-455806'
    at com.google.cloud.spark.bigquery.repackaged.com.google.api.gax.rpc.ApiExceptionFactory.createException(ApiExceptionFactory.java:41)
    at com.databricks.backend.daemon.driver.DriverWrapper.run(DriverWrapper.scala:335)
  ② Diagnose error
  Assistant Quick Fix: ON

```

After debugging the error, we figure out that, we forgot to add following permission to the service account.

1. Assign the following roles to the user or service account:

- o BigQuery Read Session User
- o BigQuery Data Viewer

[← Create service account](#)

Service account details

Grant this service account access to project (optional)

Grant this service account access to GCP Databricks Project so that it has permission to complete specific actions on the resources in your project.

[Learn more](#)

Role	BigQuery Read Session User	IAM condition (optional) ?	+ Add IAM condition
Access to create and use read sessions		Delete	
Role	BigQuery Data Viewer	IAM condition (optional) ?	+ Add IAM condition
Access to view datasets and all of their contents		Delete	
Role	BigQuery Data Editor	IAM condition (optional) ?	+ Add IAM condition
Access to edit all the contents of datasets		Delete	
Role	BigQuery Job User	IAM condition (optional) ?	+ Add IAM condition
Access to run jobs		Delete	

[+ Add another role](#)

[Continue](#)

 **databricks**

[New](#) [Workspace](#) [Recents](#) [Catalog](#) [Workflows](#) [Compute](#) [Marketplace](#)

[SQL](#) [SQL Editor](#) [Queries](#) [Dashboards](#) [Genie](#) [Alerts](#) [Query History](#) [SQL Warehouses](#)

Data Engineering [Job Runs](#) [Data Ingestion](#) [Pipelines](#)

Read Data From BigQuery Sample Data Python [star](#)

File Edit View Run Help Last edit was 19 minutes ago

1

```
table = "bigquery-public-data.samples.shakespeare"
df = spark.read.format("bigquery").option("table",table).load()
df.createOrReplaceTempView("shakespeare")
```

2

```
%sql
select * from shakespeare
```

(1) Spark Jobs

_sqldf: pyspark.sql.dataframe.DataFrame = [word: string, word_count: long ... 2 more fields]

Table +

word	word_count	corpus	corpus_date
augurs	1	sonnets	0
dimm'd	1	sonnets	0
plagues	1	sonnets	0
treason	1	sonnets	0
...	1	sonnets	0

As per my knowledge, we are not allowed to attach multiple service accounts to a single Databricks cluster. GCP IAM doesn't support "multiple active identities" on the same compute resource.

Workarounds:

- Use a Service Account with Access to Multiple Resources: Grant one service account access to all required resources (e.g., BigQuery, GCS buckets, Pub/Sub). This is the recommended pattern in most use cases.
- Assign Different Service Accounts to Each Job Cluster: If your workflow requires different identities for data isolation, break it into separate job clusters using different service accounts.

To run the pipeline end to end, we need to have two active Databricks clusters, however, I am able to keep only one cluster due to limited resource for this demo.

The screenshot shows the Databricks Compute interface. At the top, there's a navigation bar with tabs for All-purpose compute, Job compute, SQL warehouses, Pools, and Policies. Below the navigation is a search bar and filter options. A prominent button at the top right says "Create with Personal Compute".

The main table lists two clusters:

State	Name	Policy	Runtime	Active memory	Active cores	Active DBU / h	Source	Creator	Notebooks
Green	BQConnectCluster	Personal Compute	15.4	32 GB	4 cores	0.96	UI	Abhijit Das	1
Red	GCSM Cluster	-	15.4	-	-	-	UI	Abhijit Das	-

In the middle section, there's a cluster configuration form for "GCSM Cluster". It includes fields for Configuration (set to "Unrestricted"), Policy (radio button for "Single node" selected), Access mode (radio button for "Single user or group access" selected), and a dropdown for "Dedicated (formerly: Single user)" set to "Abhijit Das". A tooltip on the "GCSM Cluster" link states: "Gcp Resource Quota Exceeded: GCP Error: GCP API call failed: Quota 'SSD_TOTAL_GB' exceeded. Limit: 250.0 in region us-west4."

The "Performance" section specifies "Databricks Runtime Version" as "15.4 LTS (includes Apache Spark 3.5.0, Scala 2.12)". It also includes checkboxes for "Use Photon Acceleration" and "Enable autoscaling local storage". A "Terminate after" field is set to "15 minutes of inactivity".

The "Tags" section indicates "No custom tags" and "Automatically added tags". There's also a "Advanced options" link.

I sent a request to increase quota for SSD, which you can track:

Quotas & System Limits for project "GCP Databricks Project"									
QUOTAS & SYSTEM LIMITS		INCREASE REQUESTS							
<input type="text"/> Filter Enter property name or value									
Request number	Service	Quota	Dimensions (e.g. location)	Original value	Requested value	Approved value	Request time	Response time	Status
[REDACTED]	Compute Engine API	CPUs (all regions)		0	16	16	5/10/25, 9:52 AM	5/10/25, 9:52 AM	Approved
[REDACTED]	Compute Engine API	C2D CPUs	region: us-west4	8	16	16	5/10/25, 9:52 AM	5/10/25, 9:52 AM	Approved
[REDACTED]	Compute Engine API	C2 CPUs	region: us-west4	8	16	16	5/10/25, 9:52 AM	5/10/25, 9:52 AM	Approved

As a workaround, I will be granting one service account access to all required resources – GCS and BigQuery.

Compute > Simple form: OFF

GCP Databricks Cluster

Configuration Notebooks (5) Libraries Event log Spark UI Driver logs Metrics Apps Spark compute UI - Master

Access mode: Single user or group access

Dedicated (formerly: Single user) Abhijit Das

Performance

Databricks Runtime Version: 15.4 LTS (includes Apache Spark 3.5.0, Scala 2.12)

Use Photon Acceleration

Node type: c2-standard-4 (16 GB Memory, 4 Cores)

Enable autoscaling local storage

Terminate after 15 minutes of inactivity

Tags

No custom tags

Automatically added tags

Advanced options

Google Service Account: databricks-bigquery-demo@gcp-databricks-project-459406.iam

Instances Spark Logging Init Scripts JDBC/ODBC

Availability zone: auto # local SSDs

APIs & Services

Credentials + Create credentials Delete Restore deleted credentials

Enabled APIs & services: Create credentials to access your enabled APIs. Learn more

Remember to configure the OAuth consent screen with information about your application.

API Keys

Name	Creation date	Restrictions
No API keys to display		

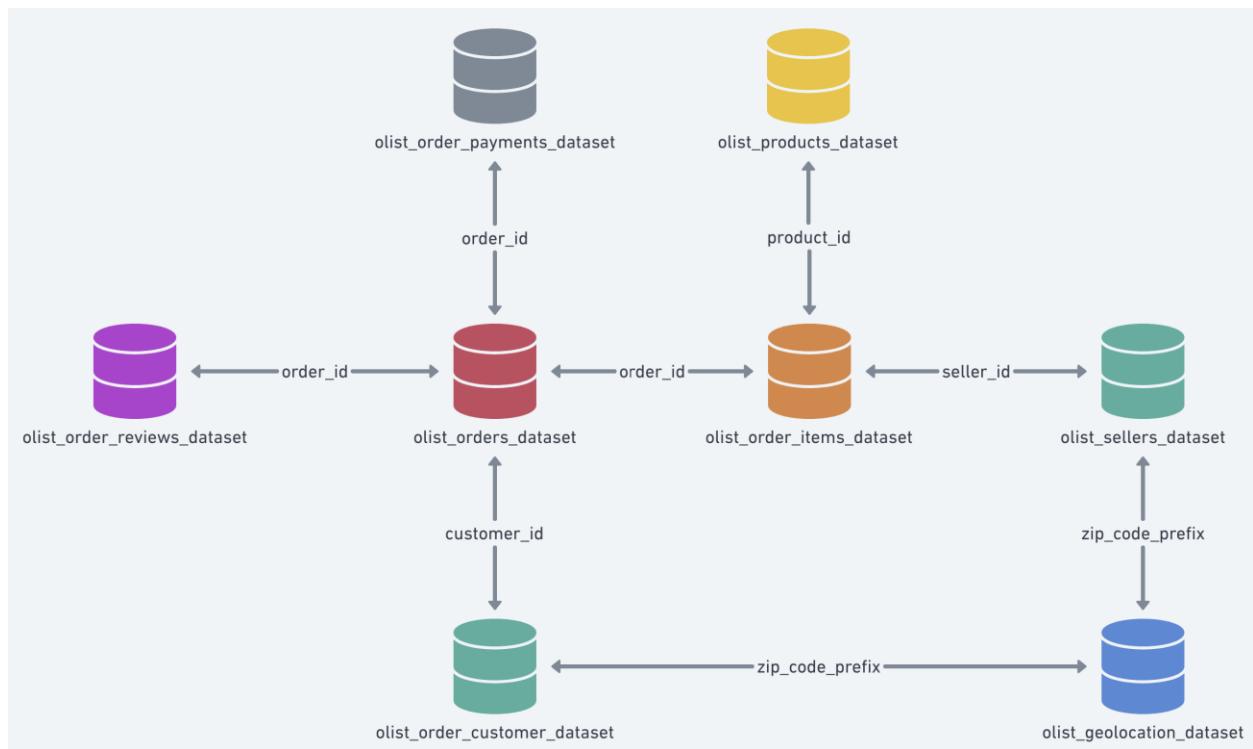
OAuth 2.0 Client IDs

Name	Creation date	Type	Client ID
No OAuth clients to display			

Service Accounts

Email	Name
[REDACTED]	Compute Engine default service account
[REDACTED]	Databricks Compute Service Account
databricks-bigquery-demo	databricks-bigquery-demo
[REDACTED]	databricks-gcp-mount

I will be using following dataset to prepare this data pipeline.



Environment Setup and common utility notebooks for bronze, silver and gold data processing and load.

Environment Setup Python ▾ Tabs: OFF ▾ ⭐

File Edit View Run Help Last edit was 15 hours ago

Run all GCP Databricks Cluster Schedule

GCS Databricks Mount

Markdown ☰ ⌂

```

▶ ✓ 4 days ago (13)
  bucket_name = "e-commerce-raw-data"
  mount_name = "gcsmount_ecommerce"
  dbutils.fs.mount(
    source = f"gs://{{bucket_name}}",
    mount_point = f"/mnt/databricks/{{mount_name}}",
    extra_configs = {"fs.gs.project.id": "gcp-databricks-project-459406"}
  )
  True

▶ ✓ 0:15 AM (1s)
  dbutils.fs.ls("/mnt/databricks/gcsmount_ecommerce/customers/olist_customers_dataset.csv")
  [FileInfo(path='/dbfs:/mnt/databricks/gcsmount_ecommerce/customers/olist_customers_dataset.csv', name='olist_customers_dataset.csv', size=9033957, modificationTime=1746942004815)]

▶ ✓ 0:16 AM (1s)
  df_csv = spark.read.format("csv").option("header", "true").load("/mnt/databricks/gcsmount_ecommerce/customers/olist_customers_dataset.csv")
  display(df_csv)
  (2) Spark Jobs
  df_csv: pyspark.sql.dataframe.DataFrame = [customer_id: string, customer_unique_id: string ... 3 more fields]

  Table + 
  ┌─────────┐ ┌─────────┐ ┌─────────┐ ┌─────────┐ ┌─────────┐
  │customer_id│customer_unique_id│customer_zip_code_prefix│customer_city│customer_state│
  └─────────┘└─────────┘└─────────┘└─────────┘└─────────┘
  1 06b899e2ba1a1fb88172c00ba8bc7 861ef4711a542e4b93843c6dd7febb0 14409 franca SP
  
```

BigQueryConnectTest Python ▾ Tabs: OFF ▾ ⭐

File Edit View Run Help Last edit was 4 days ago

Run all GCP Databricks Cluster Schedule Share

BigQuery Connect Test

Python ☰ ⌂

```

▶ ✓ 4 days ago (8)
  %bq
  !table ~ "bigquery-public-data.samples.shakespeare"
  df = spark.read.format("bigquery").option("table", table).load()
  df.createOrReplaceTempView("shakespeare")
  df
  df: pyspark.sql.dataframe.DataFrame = [word: string, word_count: long ... 2 more fields]

  Table + 
  ┌─────────┐ ┌─────────┐ ┌─────────┐ ┌─────────┐
  │word│word_count│corpus│corpus_date│
  └─────────┘└─────────┘└─────────┘└─────────┘
  1 LVII 1 sonnets 0
  2 augurs 1 sonnets 0
  3 dimm'd 1 sonnets 0
  4 plagues 1 sonnets 0
  
```

Utility Module has functions required for all three layers. Those functions will be used as a library, if we need to add more new datasets in future.

util_mod

util_mod Python Tabs: OFF 3 days ago (c1s)

```
#Bronze data loader
def load_raw_data(folders):
    from pyspark.sql.functions import lit, current_date
    dataframes = []
    load_date = current_date()
    for folder in folders:
        df_name = folder.split('/')[-2]
        df = spark.read.format("csv").option("inferSchema", "true").option("header", "true").load(folder)
        df = df.withColumn("load_date", load_date)
        df.write.mode("overwrite").partitionBy("load_date").saveAsTable(f"bronze_db.{df_name}")

    # dataframes[df_name] = df
    print(f"Data loaded successfully for {df_name} in bronze_db")
    print("All data loaded successfully")
# return dataframes
```

util_mod Python 2 3 days ago (c1s)

```
#silver data loader
def upsert_latest_partition_data(tables, target_db):
    from pyspark.sql.functions import col, row_number, current_timestamp
    from pyspark.sql.window import Window

    for table in tables:
        bronze_table = f"bronze_db.{table}"
        target_table = f"[target_db].{table}"

        # Load the latest partition data
        latest_partition = spark.sql(f"SELECT MAX(load_date) as max_date FROM {bronze_table}").collect()[0]['max_date']
        latest_data = spark.table(bronze_table).filter(col('load_date') == latest_partition)

        # Add insert_dt and update_dt columns
        latest_data = latest_data.withColumn("insert_dt", current_timestamp()).withColumn("update_dt", current_timestamp())
```

util_mod Python 2 3 days ago (c1s)

```
# Upsert using merge
deduped_data.createOrReplaceTempView("source")
merge_sql = f"""
MERGE INTO {target_table} AS target
USING source
ON {merge_condition}
WHEN MATCHED THEN
    UPDATE SET {", ".join([f"target.{col} = source.{col}" for col in deduped_data.columns if col != "update_dt"])}
WHEN NOT MATCHED THEN
    INSERT ({", ".join(deduped_data.columns)}) VALUES ({", ".join([f"source.{col}" for col in deduped_data.columns])})
"""

spark.sql(merge_sql)
print(f"Upsert completed for {table} in {target_db}")
print("All silver data loaded successfully")
```

util_mod Python 3 3 days ago (c1s)

```
#Gold data loader
def prepare_dataframes_from_tables(tables, db_name='silver_db'):
    for table in tables:
        df_name = f"[table].df"
        df = spark.table(f"[db_name].[table]").drop("insert_dt", "update_dt", "load_date")
        globals()[df_name] = df

def load_to_gold_db(df, table_name):
    df.write.format("delta").mode("overwrite").option("mergeSchema", "true").saveAsTable(f"gold_db.{table_name}")

# def load_to_bigquery(df, table_name):
#     table_name = "gcp-databricks-project-459406.ecomdata" + "." + table_name
#     df.write.format("bigquery").option("table", table_name).option("temporaryGcsBucket", "e-commerce-raw-data").save()

def load_to_bigquery(df, table_name):
    table_name = "gcp-databricks-project-459406.ecomdata" + "." + table_name
    df.write.format("bigquery").option("table", table_name).option("temporaryGcsBucket", "e-commerce-raw-data").mode("overwrite").save()
```

Bronze Layer loads data from GCS (bucket – using Databricks mount) to Databricks dronze_db as delta table.

1. BronzeDataProcess Python Tabs OFF

Raw data load to Bronze Layer - Bronze Layer Data Processing

```

folders = [
    'dbfs:/mnt/databricks/gcsmount_ecommerce/category_name/',
    'dbfs:/mnt/databricks/gcsmount_ecommerce/customers/',
    'dbfs:/mnt/databricks/gcsmount_ecommerce/geolocation/',
    'dbfs:/mnt/databricks/gcsmount_ecommerce/order_items/',
    'dbfs:/mnt/databricks/gcsmount_ecommerce/order_payments/',
    'dbfs:/mnt/databricks/gcsmount_ecommerce/order_reviews/',
    'dbfs:/mnt/databricks/gcsmount_ecommerce/orders/',
    'dbfs:/mnt/databricks/gcsmount_ecommerce/products/',
    'dbfs:/mnt/databricks/gcsmount_ecommerce/sellers/'
]

load_raw_data(folders)

```

Data loaded successfully for customers in bronze_db
Data loaded successfully for geolocation in bronze_db
Data loaded successfully for order_items in bronze_db
Data loaded successfully for order_payments in bronze_db
Data loaded successfully for order_reviews in bronze_db
Data loaded successfully for orders in bronze_db
Data loaded successfully for products in bronze_db
Data loaded successfully for sellers in bronze_db
All data loaded successfully

```

dbutils.notebook.exit("SUCCESS")

```

Silver Layer loads data from bronze_db to silver_db as delta table. It performs upsert with data cleaning, dedup and other processing. It will read only latest partition from bronze_db table.

2. SilverDataProcess Python Tabs OFF

Silver Layer Data Processing

```

tables = ['customers', 'geolocation', 'order_items', 'order_payments', 'order_reviews', 'orders', 'products', 'sellers']
target_db = 'silver_db'
upsert_latest_partition_data(tables, target_db)

```

Upsert completed for customers in silver_db
Upsert completed for geolocation in silver_db
Upsert completed for order_items in silver_db
Upsert completed for order_payments in silver_db
Upsert completed for order_reviews in silver_db
Upsert completed for orders in silver_db
Upsert completed for products in silver_db
Upsert completed for sellers in silver_db
All silver data loaded successfully

```

dbutils.notebook.exit("SUCCESS")

```

Gold layer will perform aggregation on silver_db tables and prepare final order_details table in gold_db. This order_details table (in gold layer) will be used to create / load different aggregated tables in BigQuery (consumption / serving). Power-BI reports will be generated from those tables in BigQuery.

The screenshot shows a Databricks notebook titled "3. GoldDataProcess" in Python mode. The notebook contains several code cells and a section titled "Aggregation" with the following bullet points:

- #Total Revenues Per Seller
- #Total Orders Per Customer
- #Average Review Score Per Seller
- #Most Sold Products (Top 10)
- #Top Customers By Spending

The code cells include:

```

8
# full_orders_df.count() #54,442,883 #1,042,785 - 2 mins to count without rep...
9
# full_orders_df = spark.read.table('gold_db.order_details')
load_to_bigquery(full_orders_df, 'orders.details')
# full_orders_df = pyspark.sql.DataFrame[order_id:string, customer_id:string ... 42 more fields]
10
# Total Revenues Per Seller
from pyspark.sql.functions import sum, desc
seller_revenue_df = full_orders_df.groupBy('seller_id').agg(sum('price').alias('total_revenue')).orderBy(desc('total_revenue')).limit(100)
# load_to_gold_db(seller_revenue_df, 'seller_revenue')
load_to_bigquery(seller_revenue_df, 'seller_revenue')

```

Now we have prepared a workflow (data-pipeline) to run those steps in sequence. This can be scheduled or run on demand.

The screenshot shows the Databricks Workflows interface for a pipeline named "ecommerce_order_details". The pipeline has three tasks:

- raw_data_load: Launched on May 14, duration 38m 8s
- processed_data_load: Launched on May 14, duration 19m 4s
- gold_data_load: Launched on May 14, duration 25m 32s

The pipeline status is "Succeeded". On the right, there are sections for "Job details", "Git", "Schedules & Triggers", and "Compute".

Job details

- Job ID: 503880719997047
- Creator: Abhijit Das
- Run as: Abhijit Das
- Tags: None
- Description: Add description
- Lineage: 17 upstream tables, 22 downstream tables

Git

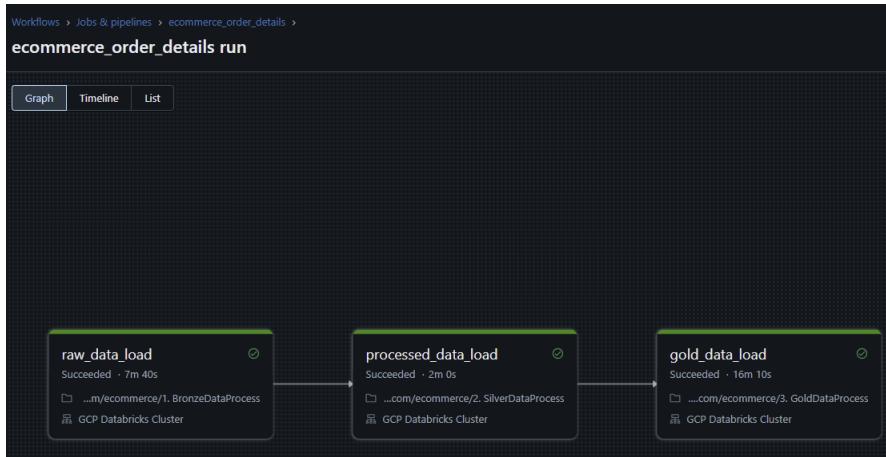
- Not configured
- Add Git settings

Schedules & Triggers

- None
- Add trigger

Compute

- GCP Databricks Cluster
- Single node: c2-standard-4 - Release: 15.4.15
- View details | Swap | Spark UI | Logs | Metrics



Workflows > Jobs & pipelines > ecommerce_order_details > Run 1077841393060873 >
gold_data_load run Succeeded

Output

Gold Layer Data Processing

```

✓ 02s
%run /Workspace/Users/ /util_mod
2

✓ 2.34s
3
tables = ['orders', 'order_items', 'sellers', 'products', 'customers', 'geolocation', 'order_reviews', 'order_payments']
prepare_dataframes_from_tables(tables)

4
customers_df: pyspark.sql.DataFrameDataFrame = [customer_id: string, customer_unique_id: string ... 3 more fields]
geolocation_df: pyspark.sql.DataFrameDataFrame = [geolocation_zip_code_prefix: integer, geolocation_lat: double ... 3 more fields]
order_items_df: pyspark.sql.DataFrameDataFrame = [order_id: string, order_item_id: integer ... 5 more fields]
order_payments_df: pyspark.sql.DataFrameDataFrame = [order_id: string, payment_sequential: integer ... 3 more fields]
order_reviews_df: pyspark.sql.DataFrameDataFrame = [review_id: string, order_id: string ... 5 more fields]
orders_df: pyspark.sql.DataFrameDataFrame = [order_id: string, customer_id: string ... 6 more fields]
products_df: pyspark.sql.DataFrameDataFrame = [product_id: string, product_category_name: string ... 7 more fields]
sellers_df: pyspark.sql.DataFrameDataFrame = [seller_id: string, seller_zip_code_prefix: integer ... 2 more fields]

```

```

✓ 0.34s
4
full_orders_df = orders_df.join(order_items_df, 'order_id', 'inner') \
    .join(products_df, 'product_id', 'inner') \
    .join(sellers_df, 'seller_id', 'inner') \
    .join(customers_df, 'customer_id', 'inner')

full_orders_df = full_orders_df.join(geolocation_df, full_orders_df.customer_zip_code_prefix == geolocation_df.geolocation_zip_code_prefix, 'left') \
    .join(order_reviews_df, 'order_id', 'left') \
    .join(order_payments_df, 'order_id', 'left')

```

Task run

Details

- Job ID: [REDACTED]
- Job run ID: [REDACTED]
- Task run ID: 534449042674481
- Run as: Abhijit Das
- Launched: Manually
- Started: May 17, 2025, 10:38 PM
- Ended: May 17, 2025, 10:55 PM
- Duration: 16m 10s
- Queue duration: -
- Status: Succeeded
- Lineage: 17 upstream tables, 22 downstream tables

Notebook

/Workspace/Users/ /ecommerce/3. GoldDataProcess

Compute

GCP Databricks Cluster

Create dataset in your region for your project-id

Create dataset

Project ID * **gcp-databricks-project-459406** Change

Dataset ID * **ecomdata**

Letters, numbers, and underscores

Location type **Region**

Region Specify a region to colocate your datasets with other Google Cloud services.

Multi-region Allow BigQuery to select a region within a group to achieve higher quota limits.

Region * **us-west4 (Las Vegas)**

External Dataset

The selected region supports the following external dataset types: Cloud Spanner

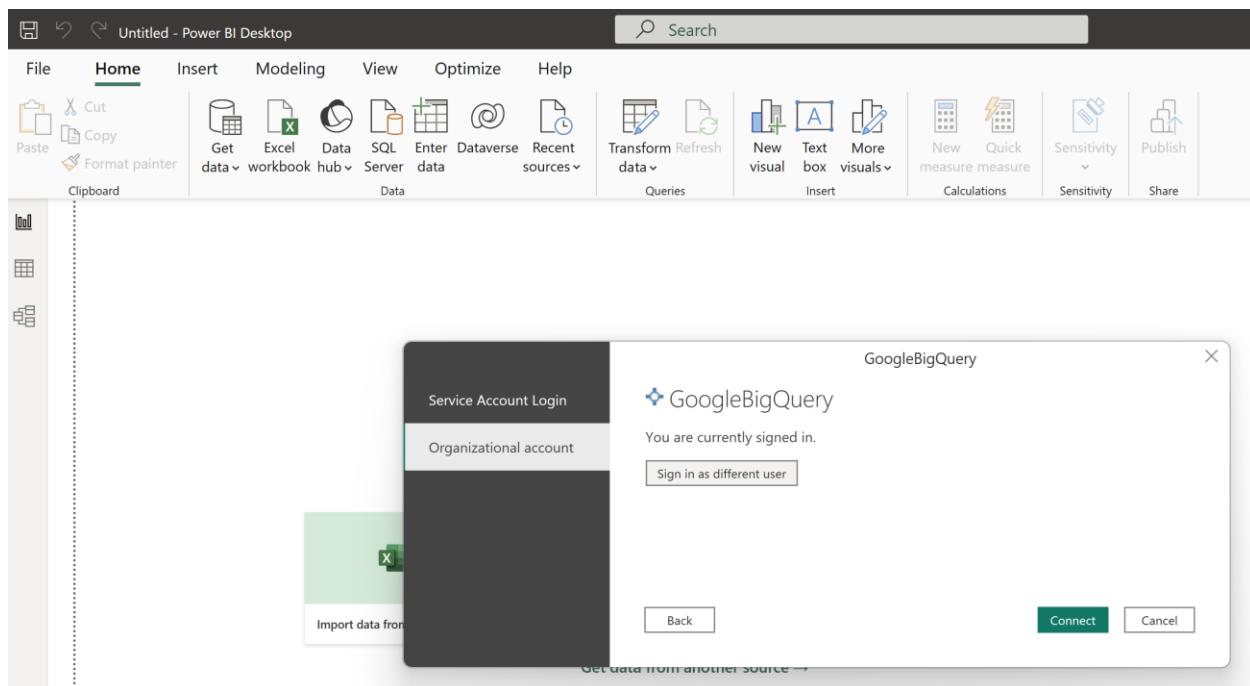
Link to an external dataset

Tags

Advanced options

Create dataset Cancel

Connect to BigQuery from Power BI



Navigator

Display Options

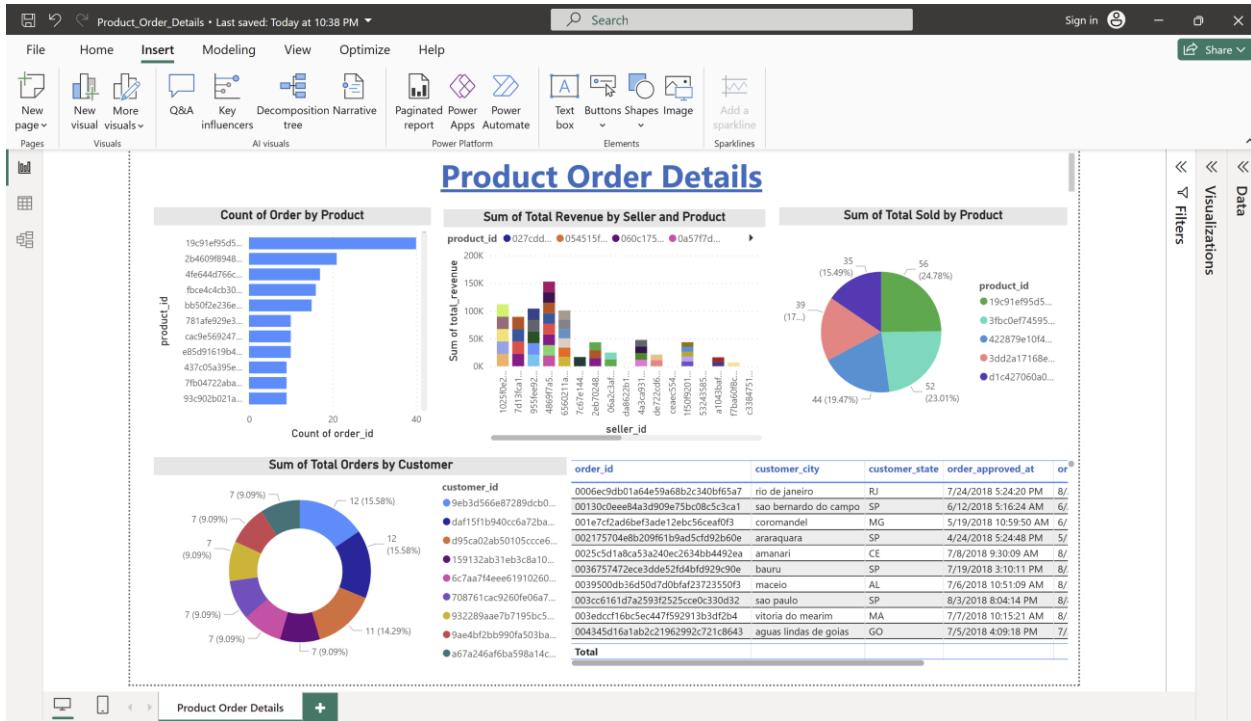
- ▲ Google BigQuery [3]
 - ▷ bigquery-public-data
 - ▷ booming-cascade [REDACTED]
 - ▲ gcp-databricks-project-[REDACTED]
 - ▲ ecomdata [2]
 - seller_revenue
 - shakespeare

seller_revenue

seller_id	total_revenue
1025f0e2d44d7041d6cf58b6550e0bfa	22378.71
7d13fca15225358621be4086e1eb0964	22288.84
955fee9216a5b617a5c0531780ce60	20814.03
fa1c13f2614d7b5c4749cbc52fecda94	19385.85
4869f7a5dfa277a7dca6462dcf3b52b2	19096.8
6560211a19b47992c3666cc44a7e94c0	16783
7c67e1448b00f6e969d365cea6b010ab	16583.41
2eb70248d66e0e3ef83659f71b244378	14410.39
7681ef142fd2c19048da7430856b5588	14007.98
966cb4760537b1404caedd472cc610a5	13912.75
b839e41795b7f3ad94cc2014a52f6796	12394
06a2c3af7b3aee5d69171b0e14f0ee87	12359.7
da8622b14eb17ae2831f4ac5b9dab84a	11878.59
4a3ca9315b744ce9f8e9374361493884	11856.6
5dceca129747e92ff8ef7a997dc4f8ca	11559.45
40db9e9aa57f7bb151bcda6b0f9bdb7	11492
de722cd6dad950a92b7d4f82673f8833	10541.6
d23019c84ffae2d5ef2270367b8605fc	9665.7
04308b1ee57b6625f47df1d56f00eedf	9307.5
ceaec5548eefc6e23e6607c5435102e7	9125.63
1f50f920176fa81dab994f9023523100	8671.48
53243585a1d6dc2643021fd1853d8905	8635.18
2bf6a2c1e71bbd29a4ad64e6d3c3629f	8229.6
c70c1b0d8ca86052f45a432a38b73958	8095.96

Select Related Tables Transform Data

Here is your Power BI Report



Now it is time to clean up resources to avoid any unwanted charges.

https://accounts.gcp.databricks.com/workspaces/account_id=741d3444-1641-4454-b5d3-fba0fbefef7

Workspaces

Name	Status	Pricing tier	Region	Created	View	Delete
gcp_databricks_demo	Running	Premium	us-west1	last Sunday at 10:02 PM	Open	Delete

Workspaces

Name	Status	Pricing tier	Region	Created	Open
gcp_databricks_demo	Running	Premium	us-west1	04/06/2025	Open

Confirm delete

Are you sure you want to delete the workspace?

Enter gcp_databricks_demo to confirm

gcp_databricks_demo

[Cancel](#) [Confirm delete](#)

The screenshot shows the Google Cloud Metastore configuration page for the catalog 'metastore_gcp_us_west1'. The configuration tab is selected. A modal dialog box titled 'Delete metastore' is open, containing a warning message: 'Warning: This action will delete all metadata and permissions associated with metastore_gcp_us_west1. This action cannot be undone. Underlying data stored in your cloud tenant is not modified.' Below the message is a text input field with the placeholder 'Type metastore_gcp_us_west1 to confirm deletion' containing the text 'metastore_gcp_us_west1'. At the bottom of the dialog are 'Cancel' and 'Delete' buttons.

The screenshot shows the Google Cloud Marketplace search results for 'databricks'. The search bar at the top contains the text 'databricks'. The results section shows 15 results, with the first item being 'Databricks - SaaS & APIs'. The listing includes a red Databricks logo, a brief description: 'Powered by Delta Lake, Databricks combines the best of data warehouses and data lakes into a lakehouse architecture, giving you one platform to collaborate on all of your data, analytics and AI workloads. Notebooks: Build data science, data engineering and machine learning notebooks using Python, SQL, R, Scala. Collaborate on these notebooks with your entire data team. Analytics: Use SQL Analytics to query and visualize dat...', and a '14 Day Free Trial' button.

☰ Google Cloud GCP Databricks Project

Product details Databricks

OVERVIEW PRICING DOCUMENTATION SUPPORT RELATED PRODUCTS

Pricing

The product was purchased on 4/3/25.

MANAGE ORDERS

Features

Here are the features supported by this Product

- Managed Apache Spark: Yes
- Optimized Delta Lake: Yes
- Cluster Autopilot: Yes
- Jobs Scheduling & Workflow: Yes
- Notebooks & Collaboration: Yes
- Databricks Runtime for ML: Yes
- Optimized Runtime Engine: Yes
- Administration Console: Yes
- Single Sign-On (SSO): Yes
- Role-based Access Control: Yes
- Token Management API: Yes

Usage fee

Usage fee is billed every month

Databricks Consumption Units	USD 1.00 /unit
------------------------------	----------------

☰ Google Cloud

Your orders for this product

VIEW BILLING REPORTS VIEW BILLING COMMITMENTS SEND FEEDBACK

Select a billing account * My Billing Account

This page contains all your orders for the Databricks product. To manage other product orders visit [Your orders](#).

Filter Filter by column name or chart value

Status	Order number	Order title	Provider	Product	Plan	Next plan	Auto-renew	Purchase date	Start date	End date	Payment schedule
Active ⓘ	e4716c...		Databricks	Databricks	Databricks	N/A	N/A	04/03/2025	04/03/2025	-	Postpay

⋮ View billing report Contact support Cancel order

≡ Google Cloud

>Your orders for this product

Select a billing account *

My Billing Account

This page contains all your orders for the Databricks product. To manage other product orders visit [Your orders](#).

Filter Filter by column name or chart value

Status	Order number	Order title	Provider	Product	Plan
Active ⓘ	e4716c...		Databricks	Databricks	Databricks

Cancel order

You can't undo this action.

By canceling the following will happen:

- Your order will be canceled immediately
- Google will notify the partner, and they may terminate your access to the applicable product and delete resources associated with this order
- Final charges for this order will be based on the product usage

Here's the order information:
Provider: Databricks
Product: Databricks

Order number:

Confirm cancellation by typing the order number below

Order number *

QUIT CANCEL ORDER